Anti-Rex Aptamers as RNA Mimics of the Rex-Binding Element

Scott Baskerville[1], Maria Zapp[2], and Andrew D. Ellington[1]

[1]Departments of Chemistry and Microbiology, Indiana University

Bloomington, IN., 47405

812-855-6071 phone

812-855-8300 fax

[2]Howard Hughes Medical Institute, Program in Molecular Medicine

University of Massachusetts Medical Center

373 Plantation Avenue, Worcester, MA 01605

ABSTRACT

RNA molecules that bind tightly and specifically to a Rex fusion protein have been isolated from a conformationally constrained pool of random sequence RNAs. These RNAs effectively mimic a number of features of the wild-type *Rex-binding element (XBE)*. The aptamers compete with the wild-type *Rex-binding element* for binding to Rex *in vitro*. The highest-affinity aptamers can specifically bind the Rex ARM, and do not recognize the functionally analogous Rev protein or its ARM. The anti-Rex aptamers can also functionally substitute for the *Rex-binding element in vivo*, a result which supports a previously proposed model for mRNA transport in which the viral genome serves as a passive platform for assembling and co-opting the cellular transport apparatus. Characteristic sequence and structural motifs found in some of the anti-Rex aptamers may provide insights into the ability of Rex to bind the *Rev-responsive element*. Overall, these studies suggest that anti-Rex aptamers may serve as effective RNA decoys of Rex.

INTRODUCTION

The human T-cell leukemia virus (HTLV-I) is etiologically associated with adult T-cell Leukemia (ATL), a malignancy of CD4+ cells (Yoshida et al., 1982; Hinuma et al., 1981), and with tropical spastic paraparesis (TSP), a degenerative neuropathy (Osame et al., 1986; Gessain et al., 1985). Estimates of infected individuals range from 10 to 20 million worldwide (Murphy et al., 1989). HTLV-I replication is largely regulated by the viral-encoded proteins Tax and Rex. Tax is a nuclear *trans*-activator that forms complexes with cellular transcription factors and stimulates transcription from promoters within the viral LTR (Beraud *et al.*, 1991; Marriott *et al.*, 1990; Zhao and Giam, 1991). Rex functions *post*-transcriptionally, enhancing the cytoplasmic appearance of incompletely spliced mRNAs encoding HTLV-I structural proteins and suppressing the transport of mRNAs encoding of regulatory proteins such as Tax and Rex (Gröne *et al.*, 1996; Hanly *et al.*, 1989; Hidaka *et al.*, 1988; Inoue *et al.*, 1987). In effect, Rex modulates the switch between the early phase of the life cycle, in which viral regulatory proteins are produced, and the late phase, in which viral particles are assembled (reviewed in Cullen, 1992). Because of its central role in regulating viral replication and the progression of infection, Rex is an excellent target for the development of anti-viral compounds.

The appearance of singly or unspliced messages in the cytoplasm is mediated by the *trans*-acting Rex protein (Inoue et al., 1987; Hidaka et al., 1988) and the *cis*-acting *Rex-responsive element* (*XRE*) (Seiki et al., 1988; Hanly et al., 1989; Ahmed et al., 1990). Mutational analysis has indicated that an arginine-rich motif (ARM) at the amino terminus of Rex contributes to RNA-binding (Hammes et al., 1993; Grassmann et al., 1991; Bogerd et al., 1991). The sequences and structures that contribute to the function of the *XRE* have been studied by a variety of methods, including site-directed mutation, *in vitro* selection, chemical protection, and modification interference (Baskerville *et al.*, 1995; Gröne *et al.*, 1994; Bogerd *et al.* 1992; Toyoshima *et al.*, 1990; Hanly *et al.*, 1989). Overall, these

experiments have discerned that the primary *Rex-binding element* (*XBE*) maps to a stem-bulge structure that spans residues 169-180 and 200-211 of the *XRE* (Figure 1a).

Since Rex interacts with a defined RNA substrate, it may prove possible to treat or prevent HTLV-I infections by using small RNA molecules to decoy Rex. For example, the HIV-1 *TAR* element and the HIV-1 *Rev-binding element* (*RBE*) are small defined RNA molecules that bind to the HIV-1 Tat and Rev proteins, respectively. When these RNA molecules are expressed in cells, they decoy the function of Tat and Rev, and inhibit HIV-1 infection (Lee et al., 1994; Sullenger et al., 1990).

In vitro selection has previously proven to be an invaluable tool for identifying novel inhibitors of protein function (Ellington and Conrad, 1995; Gold, 1995; Klug and Famulok, 1994). In particular, we and others have selected RNA molecules (aptamers) from random sequence pools that bind with high affinity and specificity to the Rev protein and disrupt interactions between Rev and the *RBE* (Giver *et al.*,1993; Tuerk and MacDougal-Waugh, 1993; Bartel *et al.*,1991). When expressed in cells the anti-Rev aptamers can suppress Rev function and inhibit viral replication (Good et al., 1997).

We have now used an *in vitro* selection approach to identify novel, high-affinity RNA aptamers that can bind to the Rex protein and the Rex ARM. In addition to representing reagents that can potentially be used to inhibit HTLV-I replication, the sequences and structures of the aptamers represent a comprehensive map of the sequences and structures that can potentially be recognized by Rex. For example, it is known that the Rex protein binds to a portion of the *Rev-responsive element* and can functionally substitute for Rev (Bogerd et al., 1991; Rimsky et al., 1988). By comparing the anti-Rex aptamers with the *RRE* it may be possible to discern whether this cross-recognition is fortuitous, or is the result of a previously unknown overlap in the binding specificities of Rex and Rev.

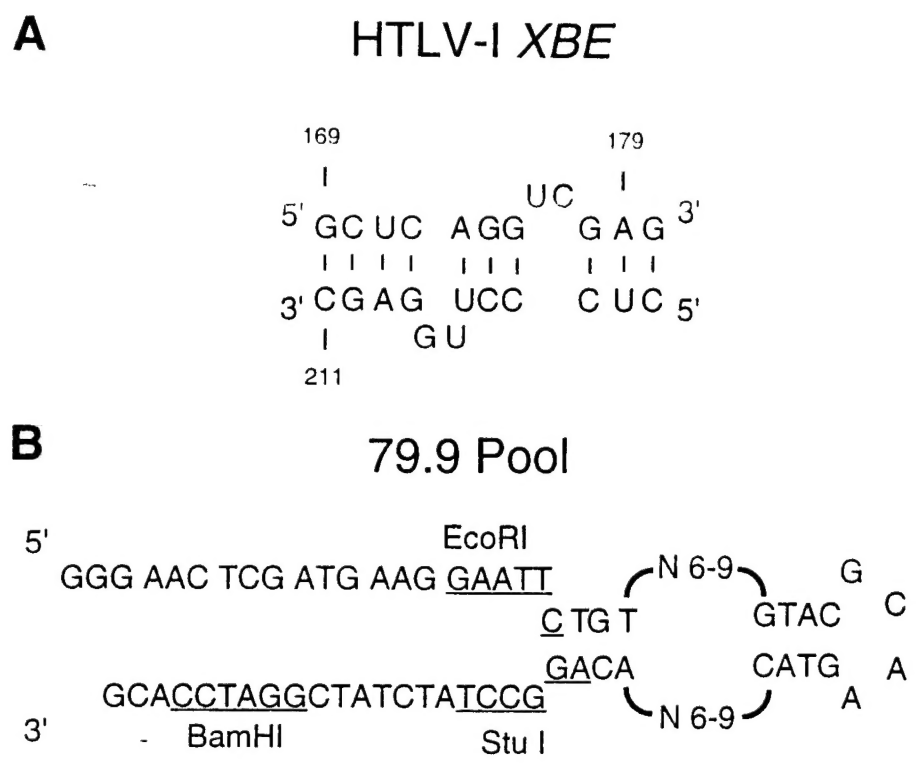


Figure 1. The *XBE* and the 79.9 Pool

(**A**) Deletion mapping (Ahmed *et al.*, 1990), mutational analysis (Gröne *et al.*,1994), modification interference (Bogerd *et al.*,1992) and *in vitro* selection experiments (Baskerville et al., 1995) have all localized the essential Rex binding element (*XBE*) to a subdomain of stem IID spanning nucleotides 169-180 and 200-211. The core of the *XBE* spans nucleotides 174-180 and 200-204 (Bogerd et al., 1992).

(**B**) The 79.9 pool contains a randomized region 12-18 nucleotides in length, flanked by constant regions used for reverse transcription and PCR amplification. A GNRA tetra-loop caps the 3' end of the engineered stem contained in the pool. The randomized region of this pool is large enough to span the core and flanking regions of the *XBE*.

Finally, the anti-Rex aptamers can potentially functionally substitute for the *XBE* and better define the role of the *XBE* in Rex-responsiveness. For example, anti-Rev aptamers have been substituted into the *RRE* in place of the *RBE* and support Rev-responsiveness(Symensma et al., 1996). These results suggested that the sequence and structure of the *RBE* played little role in Rev-responsiveness beyond binding Rev. Similarly, it is currently unclear whether the sequence and structure of the *XBE* play a role in Rex-responsiveness beyond merely providing a binding site for Rex. To address this question, the anti-Rex aptamers were cloned into the *XRE* in place of the *XBE*, and the resultant constructs were assayed for Rex-responsiveness.

MATERIALS AND METHODS

Materials

Oligonucleotide primers were synthesized on a Model 391 Applied Biosystems DNA synthesizer and purified by denaturing polyacrylamide gel electrophoresis prior to the amplification step. Dulbecco's Modified Eagle media (DMEM) and fetal bovine serum (FBS) was purchased from Gibco/BRL. The plasmid pCR II was purchased from Invitrogen. Acetyl CoA used in the CAT assays was purchased from Pharmacia; 14-C-chloramphenicol (CAT ASSAY grade) was obtained from New England Nuclear. All restriction and modifying enzymes were from Promega; Taq polymerase (AmpliTaq) by Perkin Elmer was used for the PCR amplification of aptamer DNAs. The GST-Rex fusion used in these studies has been previously described (Baskerville et al., 1995) and includes the first 79 amino acids of the Rex protein plus a stop codon at position 80.

Random Sequence RNA Pools

A completely randomized sequence pool of RNA molecules centered on a stem internal-loop-stem structure was used as a starting point for *in vitro* selection (described in Giver *et al.*, 1993) (Figure 1b). A chemically synthesized DNA oligonucleotide was used to generate the randomized sequence RNA pool. This oligonucleotide contained constant sequences flanking an internal loop varying in size from 12 - 18 nucleotides in length. Approximately 0.5 µg of the amplified template was added to an Ampliscribe (Epicentre Technologies, Madison, WI) transcription reaction. Roughly 30 µg of the initial RNA pool was obtained following purification on a 10% denaturing polyacrylamide gel.

\Specific Competitor RNA

Stem IID of the *Rex- Responsive element* was transcribed from XhoI-cut pGEM-*XRE*. This plasmid contains residues 143 through 223 of the *XRE* (numbering as in Bogerd et al., 1992); in addition to these residues, some polylinker sequences were also transcribed.

Plasmids

pGEM REM 8 is the parent vector of the aptamer/wild-type chimeric elements, an *XRE* deletion mutant previously described by Ahmed, et al., 1990 and kindly provided by W. Greene. This deletion mutant leads to a total loss of Rex responsiveness (Ahmed, et al., 1990). This plasmid contains a unique Bgl II site in place of a portion of stem IID within the *XRE*, as well as an Xba I site at each end of the *XRE* (Hanly, et al., 1989).

XRE Chimeras

pGEM REM 8 was linearized with Bgl II and ligated to the 8-5 or 9A aptamer DNA insert. Each aptamer insert contained a Bgl II site at its 5' and 3' ends, as well as an Xho I site internal to the 3' Bgl II site. The aptamer inserts were prepared by PCR amplification using the plasmid pCRII 8-5 or pCRII 9A DNA as the reaction template, and the appropriate oligonucleotide primers. The *XRE* chimeric elements were excised from the pGEM background by restriction with Xba I. These elements were then blunt end ligated into the unique Cla I site of pCMV138 (previously described in Luo et al., 1994); plasmid is identical to pDM138 (Hope et al., 1990) except that the promoter region was changed from the simian virus 40 promoter to the cytomegalovirus promoter.

Primers

Forward 8-5 primer: 5' GGG**AGATCT**CGAGATAGGCCGGCGC 3'

Reverse 8-5 primer: 5' GGG**AGATCT**GTAGGCGACGGTA 3'

Forward 9A primer: 5' GGG**AGATCT**CGAGCCTGTCCGGTGA 3'

Reverse 9A primer: 5' GGG**AGATCT**GTCGACGGGTACG 3'

Sequences underlined indicate the 5' or 3' ends of the aptamer incorporated into the *XRE* background. Sequences in bold indicate the addition of a Bgl II site (at 5' end) or a Bgl II/Xho I site (at 3' end) by PCR amplification. Construct sequences were verified by dideoxy sequencing.

Transfection Experiments

CV-1 cells were grown in 1X Delbecco's Modified Eagle Media (DMEM) containing 11% Δ FCS, 35 mM Na_2CO_3 and 25 mM Hepes, pH 7.1. Twenty four hours prior to transfection, cells were seeded at 75% confluency in 60 mM culture dishes. Transient transfection assays for Rex function were performed using a calcium phosphate-mediated DNA transfection protocol described by Lin and Green (1989) with the modifications described below. Each transfection mixture contained 1 μg of a given CMV-promoter driven *XRE* CAT reporter plasmid, 0.75 μg pcREX expression plasmid, and 2 μg of a β-galactosidase expression plasmid (pRSV-β-gal; Promega) as an internal control for both transfection efficiency and gene expression. Total DNA was adjusted to 10 μg with an empty expression plasmid containing the CMV promoter or empty pUC119 DNA. Transfected cells were harvested forty eight hours post-transfection and processed using a

CAT assay previously described by Hope, et al., 1991 and the β-gal assay of Lin and Green, 1989.

Selection Conditions

In each cycle, tRNA (Boehringer Mannheim Indianapolis, IN) and Rex protein were incubated together for 10' at ambient temperature in 1 X Binding Buffer (50 mM Tris-Cl, pH 8.0; 50 mM KCl; 30 µl). Pool RNAs in Binding Buffer (20 µl) were separately heated to 90° C for 2', cooled to ambient temperature over 10' (to equilibrate conformers), in rounds that were pre-filtered pool RNA was passed over HAWP 25 modified cellulose filters (Millipore, New Bedford, MA) to remove filter-binding sequences, and then added to the protein mixture. Specific competitor RNAs were treated similarly excluding the prefiltration step. The final reaction mix (60 µl) was left an additional 60'. The amounts of tRNA, Rex, and pool were varied in subsequent cycles as detailed in Table 1. RNA:protein complexes were separated from free RNA by vacuum filtration (5" Hg) over modified cellulose filters. Following application of the RNA:protein mixtures, filters were washed twice with 500 µl of Binding Buffer except for the final two rounds which included a dilution step in the binding reaction. Bound RNAs were eluted from filters with 400 µl 2x PK Buffer (0.2 M Tris-Cl, pH 7.6; 2.5 mM EDTA; 0.3 M NaCl; 2% SDS) at 75° C. The eluate was extracted with phenol:chloroform, precipitated with ethanol, and the pellet resuspended in 25 µl of water.

Amplification

RNA for subsequent cycles was synthesized by a combination of reverse transcription, PCR amplification, and *in vitro* transcription.

79.9 Selection Conditions

Round	[tRNA]	[79.9 Pool]	Rex (pm)	[stem IID]	Pre-Filtration	Dilution
1	1.3 µM	1.3 µM	1.8	-	3 X	-
2	6.6 µM	"	"	-	"	-
3	"	"	"	.87 µM	"	-
4	"	"	"	2.2 µM	-	-
5	"	"	.9	"	3 X	-
6	"	"	"	3.4 µM	"	-
7	"	"	"	4.2 µM	"	Yes[1]
8	13 µM	"	"	8.6 µM	"	Yes[2]

[1]Diluted to a final volume of 1 mL in 1 X Binding Buffer. This reaction was allowed to equilibrate an additional 10 minutes and then filtered.

[2]Diluted to 800 µL in 1 X Binding Buffer, then an additional 200 µl of *XRE* containing solution was added for a final concentration of 6.9 µM. This reaction was allowed to equilibrate an additional 10 minutes and then filtered.

Table 1. Selection Conditions for the 79.9 Pool.

The concentrations and amounts of nonspecific competitor tRNA, Rex fusion protein, and specific competitor (stem IID) used in each round of selection are shown. Pre-filtration of the pool prevents the accumulation of sequences that can interact with the nitrocellulose filters that were used in this selection scheme.

A portion of the extracted RNA (10 µl) was reverse transcribed in a reaction mix (20 µl) that contained 40 mM KCl; 50 mM Tris-Cl, pH 8.0; 6 mM $MgCl_2$; 0.4 mM dNTPs, 2.5 µM primer 20.86 5' CGTGGATCCGATAGATAGGC, and 5 U AMV reverse transcriptase (Seikagaku, St. Petersberg, FL). Nucleotides, DTT and enzyme were only added after an initial annealing step (3' at 75° C; 5' at ambient temperature). The reaction mix was then incubated at 42° C for 45'. For PCR amplification, 15 µl of this reaction was diluted into 85 µl of PCR reaction mix (10 mM Tris-Cl, pH 8.3; 50 mM KCl; 1.5 mM $MgCl_2$; 5% acetamide; 0.05% NP-40; 200 µM dNTPs; 2.5 U Taq polymerase; 0.5 µM 37.17 5' GGTAATACGACTCACTATAGGGAACTCGATGAAGCGA and 20.86). The reaction mix was cycled (94° C, 45"; 45° C, 1'; 72° C, 2') for as many cycles as needed to produce a product band of the correct size. RNA was synthesized from the PCR template in an Ampliscribe transcription reaction mix (Epicentre Technologies, Madison, WI) and purified on a denaturing polyacrylamide gel for subsequent cycles of selection.

Cloning and Sequencing

PCR products from the fourth and final cycle of selection (1 µg) were directly ligated into a TA cloning vector (Invitrogen, San Diego, CA) according to the protocol provided. Clones were sequenced using standard protocols for dideoxy incorporation.

Mobility Shift Assays

Following thermal equilibration in 1 X Binding Buffer (75° C, 2'; ambient temperature, 10'), ~10 ng of end labeled aptamer RNA was mixed with an increasing amount of Rex fusion protein and 2 µg of tRNA in a 10 µl volume of 1 X Binding Buffer. This reaction

mixture was allowed to equilibrate 1.5 hours at ambient temperature. Immediately prior to loading the reaction was gently mixed with 2 µl of 6X Non-Denaturing Dye (0.25% Bromophenol Blue; 40% glycerine in ddH_2O). The complexes were separated on a nondenaturing polyacrylamide gel (8% (19:1), 1X TBE, 5 Watts) and allowed to run at ambient temperature until the Bromophenol Blue had run 12 cm into the gel.

RNA Titration

When RNA was titrated against a constant amount of protein, ~30 ng of Rex protein was combined with 1 µg of tRNA in 1 X Binding Buffer (50 mM Tris-HCl pH 8.0; 50 mM KCl) and allowed to equilibrate 10 minutes at ambient temperature. Then this reaction mix was combined with varying concentrations of thermally equilibrated aptamer and (75° C 3'; ambient temperature 10') and allowed to equilibrate 1 hour at ambient temperature. To determine the specific activity of the RNA 5 µl of a 300 nM solution was spotted onto nitrocellulose and quantitated on a Phosphorimager (Molecular Dynamics, Sunnyvale CA). Moles of RNA bound were calculated from counts retained on nitrocellulose filters and plotted against the concentration of RNA in each reaction mix.

Peptide Mobility Shift

Min -apt was transcribed from annealed DNA oligonucleotides

T7 Primer 5' GAAATTAATACGACTCACTATAG 3'

Template 5' GGGCGTACCGTCGTACTTGCGTACCGGCGCCCTATAGTGA GTCGTATTAATT 3'

This sequence spans the identified core of the aptamer and is similar to the sequence of clone 39b. Products of *in vitro* transcription were purified on a 20 % denaturing

polyacrylamide gel. End labeled RNA was thermally equilibrated (50 mM KCl; 50 mM Tris-HCl, pH 8.0; 90° C 2'; cool ambient temperature). Peptides spanning the ARM of Rex and Rev were titrated against a constant amount of end labeled aptamer (Rex peptide, MPKTRRRPRRSQRKRP-am (Xu and Ellington, 1996), Rev peptide, Suc-TRQARRNRRRRWRERQRAAAAR-am (Tan et al., 1993)) (50 mM Tris-HCl; 50 mM KCl; 1 mM $MgCl_2$; 5% glycerol; ~ 5 ng RNA; 10 µl total volume) was incubated with peptide at 4° C for 1 Hour. The complexes were separated on a 10% native polyacrylamide gel (59:1; 0.5 X TBE; 350 V) run at 4° C.

Pool Assay

Pool RNAs were assayed for their ability to be co-retained with the Rex protein on modified cellulose filters. Unselected and selected RNA pools were internally labeled by including 1 µl of α-^{32}P UTP (3000 Ci / mmol, Dupont NEN, Boston, MA) in an Ampliscribe *in vitro* transcription reaction mix. RNA transcripts were gel purified and resuspended in ddH_20. Following thermal equilibration (50 mM KCl; 50 mM Tris-HCl, pH 8.0; 75° C 2'; ambient temperature, 10'), pool RNAs (0.96 µM final) were added to a mixture of tRNA (9.6 µM final), and Rex protein (~30ng) for a final volume of 40 µl and the binding reaction was allowed to equilibrate at ambient temperature for 1 hour. The mixture was then filtered over HAWP 25 modified cellulose (Millipore, New Bedford, MA) and the filters washed twice with 500 µl 1 X Binding Buffer. The amount of radioactive RNA that was co-retained with protein was quantitated using a Phosphorimager. Background binding to the filter alone was independently determined and subtracted from the protein-dependent signal.

Pool Competition

In pool competition experiments pool RNAs (20 µl) were thermally equilibrated separately and were combined and incubated with stem IID of the *XRE* (20 µl), an excess of tRNA (20 µl) and ~30 ng of Rex protein in 1 X Binding Buffer (0.59µM pool RNAs; 0.64 µM stem IID; 6.6 µM tRNA; final volume 60 µl). The reaction was incubated, filtered, and eluted as in the selection experiments. After precipitation, samples were resuspended in 4 µl stop dye (7 M urea, 1x TBE, 0.1% Bromphenol Blue) and electrophoretically separated on a denaturing 10% polyacrylamide gel. The amount of radioactivity in each band was determined using a Phosphorimager. Binding ratios were calculated using the formula [((# counts filtered)$_{pool}$ - background) / ((# counts unfiltered)$_{pool}$ - background)] / [((# counts filtered)$_{stem\ IID}$ - background) / ((# counts unfiltered)$_{stem\ IID}$ - background)]. Clones were assayed in a similar fashion. RNA was transcribed from PCR amplified template and products of *in vitro* transcription were gel purified on a 10% denaturing polyacrylamide gel. Clones and stem IID of the *XRE* were separately thermally equilibrated in 1 X Binding Buffer (90° C, 2'; cool ambient temperature 10'). tRNA and Rex fusion protein were mixed together and allowed to equilibrate for 10 minutes. An excess of thermally equilibrated clone and stem IID and were combined and equilibrated with Rex and tRNA at ambient temperature for 1 hour (0.9 µM or .45 µM stem IID; 0.64 µM clone; 0.6 µM tRNA, ~30 ng Rex) this mixture was filtered over nitrocellulose and binding ratios calculated as above. These numbers represent a rough estimate of the relative affinity of the protein for each aptamer tested. In some cases additional stem IID was added to the reaction mix to increase the signal during quantitation. This did not have any effect on the final binding ratios.

RESULTS

In vitro Selection of High Affinity RNA Elements that Interact with the Rex protein

A conformationally constrained RNA pool (79.9) that contained two random sequence tracts of 12-18 residues in length was used as a starting point for the selection of anti-Rex aptamers (Figure 1b). The 79.9 pool is similar in size and structural complexity to the wild-type *XBE* (Figure 1a), and thus RNA molecules selected from this pool should in principle be able to satisfy the same interactions with the Rex protein (Bogerd et al., 1992). Moreover, since the 79.9 pool was previously used for the selection of aptamers that could bind to HIV-1 Rev (Giver *et al.*, 1993), it should be possible to directly compare the sequences, structures, and specificities of anti-Rex and anti-Rev aptamers.

Complexes between Rex and members of the RNA pool were isolated by co-immobilization on nitrocellulose filters. The captured RNAs were eluted and amplified by reverse transcription, PCR, and *in vitro* transcription. In order to select the highest affinity binding species, a selection scheme was employed that became progressively more stringent at each round (Table 1). Transfer RNA was included in the binding reaction as a non-specific competitor and its concentration was increased through the course of the selection. stem IID of the *XRE* (containing the *XBE*) was introduced as a specific competitor in the third round and its concentration was also progressively increased. The concentration of the Rex fusion protein target was decreased either by adding less target to the binding reaction, or by diluting the binding reaction prior to capturing Rex:RNA complexes by filtration.

The selected population was assayed for its ability to bind Rex after four and eight rounds of selection and amplification. In the presence of non-specific (tRNA) and specific (*XBE*) competitors only 1.3% of the naive RNA population could bind to Rex, while 7% of the population from the fourth round of selection could bind Rex, and 13% from the

eighth round of selection. The RNA populations from rounds 0, 4, and 8 were also assayed in direct competition with the wild-type binding element (Figure 2). Labeled pool and *XBE* RNAs were mixed with limiting amounts of Rex. The binding reactions were allowed to come to equilibrium over the course of an hour, filtered over nitrocellulose, and bound nucleic acids were eluted and precipitated. Since RNA was in excess over target, different species should be retained in proportion to their relative affinities for Rex. The binding ratios for pool to wild-type RNAs were determined following electrophoretic separation and quantitation with a Phosphorimager (Molecular Dynamics, Sunnyvale CA). While unselected RNAs had a binding ratio of 0.39, the RNA population from round 4 had a binding ratio of 2.9 and the population from round 8 had a binding ratio of 8.5 (Figure 2); in other words, by the eighth cycle the selected RNA population as a whole could bind Rex almost 9-fold better than the wild-type *XBE*. The increase in binding ratio from round 4 to round 8 suggested that the pool had been winnowed until only the best binding species were retained.

Novel, High-affinity RNA Ligands for Rex

Individuals from the fourth and eighth rounds of selection were cloned and sequenced, and sequences were aligned. The round 4 aptamers could be divided into three novel families based on their primary sequences (Figure 3). The relative binding affinities of the round 4 aptamers were determined in competition with stem IID (Figure 3), and ranged from 0.6-fold of wild-type to roughly 9-fold better than wild-type. The highest affinity aptamers could be found in Class I, while aptamers of intermediate affinity were dispersed between all three families.

Two of the three families could be folded to form a consensus secondary structure. The secondary structures of individual aptamers were predicted using the program Mulfold (Zucker, 1989; Jaeger *et al.*, 1989a and 1989b).

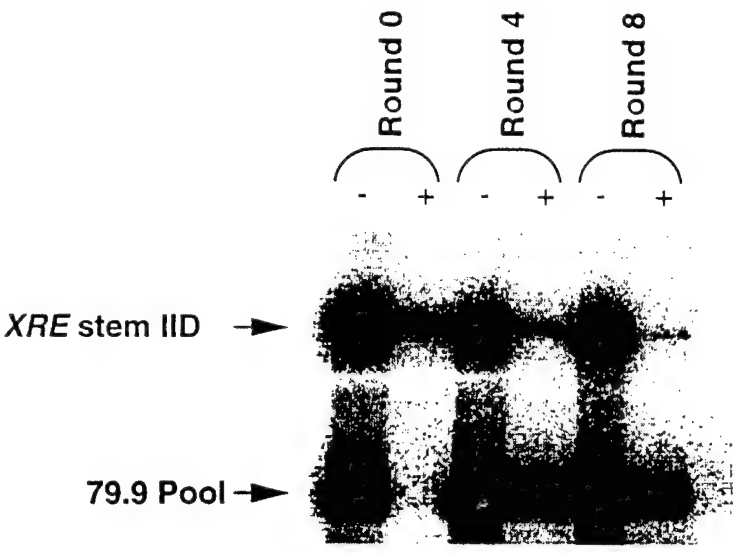


Figure 2. Pool competition assay with stem IID of the HTLV-I Rex-responsive element.

In vitro transcribed internally labeled transcripts from Round 0 (unselected), Round 4, and Round 8 were incubated separately with body labeled stem IID of the *XRE*. After one hour at ambient temperature the reaction mix was filtered over modified nitrocellulose, washed with 1 X Binding Buffer and eluted from the filters. The precipitate of the eluate (+) and an aliquot of the unfiltered reaction mix (-) was run out on a 10 % denaturing polyacrylamide gel and the products visualized on a Phosphorimager. The population of Pool 0 molecules do not effectively compete with stem IID, after multiple rounds of *in vitro* selection the population of molecules from the fourth and eighth rounds can begin to compete with stem IID for binding, illustrated by the increase in observed binding ratio; round 0 (.39), round 4 (2.9), round 8 (8.5) calculated as described in Materials and Methods.

Round 4 Aptamers

CLASS I

						Activity vs. stem II
9A	ctgt	CGACGG	gtacgcaagtac	GTCACCGG	acaggcc	0.7
13D	ctgt	CTACGG		AAATCCGG	acaggcc	0.9
20D	ctgt	ACGACG		AGCGCCGGT	acaggcc	2.7
15D	ctgt	TGCGACGGT	-	CTCCGGCT	acaggcc	3.1
1D	ctgt	GCGACGG		ACTCCGGC	aca-gcc	4.7
[illegible]						
16	ctg**g**	TCGCCCG	gtacgcaagtac	GACGGTTCG	acaggcc	1.3
6D	**t**tgt	CACATCCG		GACGGG	acaggcc	1.9
22A	ctgt	ATCCGG		CGACGGTG	acaggcc	1.9
14D	ctgt	ATCCGG		GACGGTGG	acaggcc	3.0
38B	**t**tgt	CTGCCG		TACGGTT	acaggcc	3.8
24D	ctgt	CACTTCCGG		AACGGG	acaggcc	4.7
35B	ctgt	GCTTGCCGG		CAACGGA	acaggcc	5.2
25D	ctgt	CGGCCGG		AACGGT	ac**c**ggcc	7.3±.71
5A	ctgt	AGGCCTCCG		GACGGTGC	acaggcc	8.0
15A	ctgt	CCGGG **t**		CAACGGTGA	acaggcc	8.1±1.3
*39B	ctgt	AGCCTGCCG		GACGGT	acaggcc	9.3±1.8

CLASS II

45	ctgtGGTTGAGCC		GTTGTCCT	acaggcc	1.1
17D	ctgt GTGAGCTC		TTGAGATCC	acaggcc	2.3
27	ctgt TGGCCA		TGTTGAGC	acaggcc	2.5
9D	ctgt AAGGTGCTC		TTGAGCTC	acaggcc	4.6

CLASS III

7D	ctgt	GGACGT	gtacgcaagtac	TTACCT	acaggcc	0.6
47	ctgt	GTTGATGTT		ATTCCCA	acaggcc	1.1
44	ctgt	TGATGTTT		GATTTCCG	acaggcc	1.7
19D	ctgt	GATGTT		ATTTCC	acaggcc	1.8
46	ctgt	GATGTT		ATTCCC	acaggcc	2.1
21D	ctgt	CTAGATGT		TTTCCT	acaggcc	2.3
37B	ctgtGAGCGATGT	**a**		TTTCCGCA	acag-cc	4.1

Figure 3. Sequences from round 4 and round 8 of the 79.9 selection.

Sequences from Round 4 and Round 8 were aligned based on similarities found in their primary sequence. After 4 rounds of selection and amplification the sequences could be separated into three distinct classes. By the final round of selection the Class III motif had been completely selected against, leaving the dominant Class I motif and Class II molecules. Completely conserved nucleotides are underlined, mutations in constant regions are highlighted in **bold**, and random regions contained within the pool are in CAPITAL LETTERS. Asterix (*) indicate those clones that were titrated with Rex fusion protein. All activities are expressed as the average of at least two trials.

Round 8 Aptamers

CLASS I

8-22	ctgt	CGACGGTGC	gtacgcaagtac	TCCGG	**g**caggcc
8-28,3	ctgt	CAACGGTG	--	AGCGCCGG	acaggcc
*8-5,2,26	ctgt	AGGCGACG		TCTTGCGCC	---ggcc
8-24	ctgt	AGGCGACG		GCTTGCGCC	---ggcc
*8-6,4	ctgt	GGTGCGGT	**t**	GTCCCCC	acaggcc
8-1	ctgt	CGTGCGGT	**t**	GTCCCCG	acaggcc
8-19	**t**tgt		**t**	TCCGGCTTC	acaggcc
*8-29,15	ctgt	CCGG	gtacgcaagtac	CAACGGTGA	acaggcc
*8-23	ctgt	CCGG	**t**	CAACGGTGA	acaggcc
8-18	ctgt	GCTCCGG		CAACGGTG	a-aggcc
8-13	ctgt	CTGCCG		AACGGTC	acaggcc
8-16	ctgt	GCCTGCCGG		AACGGTT	acaggcc
8-9	ctgt	AAACTTCCG		GACGGT	acaggcc
8-17	**t**tgt	CTTGTCCG		GACGGTGA	acaggcc
8-11	ctgt	ACCGG	**t**	GACGGTTTG	acaggcc
8-20	c-gt	AGCCG		GACGGTACT	acaggcc

CLASS II

8-21	ctgt	GTTGAGCC	gtacgcaagtac	GTGCC	acaggcc
8-7	ctgt	GTTGAGCAT		ATTGGCC	acaggc**t**
8-27	ctgt	TGAGCTC	-	TTGAGACT	**c**caggcc

The most abundant family of aptamers, Class I, was predicted to form a stem structure containing a centrally paired 5' CGG ... CCG 3' motif (Figure 4a). The sequence and predicted structure of this aptamer did not resemble the *XBE*. Interestingly, variants were isolated in which this stem could be oriented either towards or away from the constant sequence stem-loop. A highly conserved adenosine is found 5' to the centrally-paired region in many Class I clones (30/32) while a uridine is located 3' to this structure in a majority of the clones (22/32). These residues were predicted to participate in bulges that varied in size from 2-5 nucleotides. Class II molecules were predicted to contain a 5' GAG ... CUC 3' pairing flanked by a bulge loop containing two uridine residues (Figure 4b). This sequence and structural motif is reminiscent of an arginine-binding pocket that is found in HIV-1 *TAR* (Puglisi et al., 1992) and that has been predicted to exist in the *XBE* (Baskerville et al., 1995). Class III molecules were characterized by a conserved consensus sequence 5' NGAUG U(1-3) ... AUUCCCN 3' that could not be readily folded into an obvious, stable secondary structure.

By round 8 only Class I and II molecules remained in the population, with Class I molecules predominating. In addition, the diversity of the Class I molecules observed in round 4 had narrowed significantly, indicating that the population had been further pared to an optimal subset of high affinity structures. Both orientations of the Class I motif were still present in the population. A number of clones from the final population (Figure 3) were assayed for their ability to bind Rex (data not shown). Of these, clone 8-5 demonstrated the highest affinity for the Rex fusion protein. In addition, by comparing the sequences and predicted secondary structures of Class I ligands isolated in rounds 4 and 8, it proved possible to identify a consensus anti-Rex aptamer (Figure 4a); the consensus aptamer was similar to 8-5. Therefore, aptamer 8-5 was chosen for further characterization and development as a potential mimic of the *XBE*.

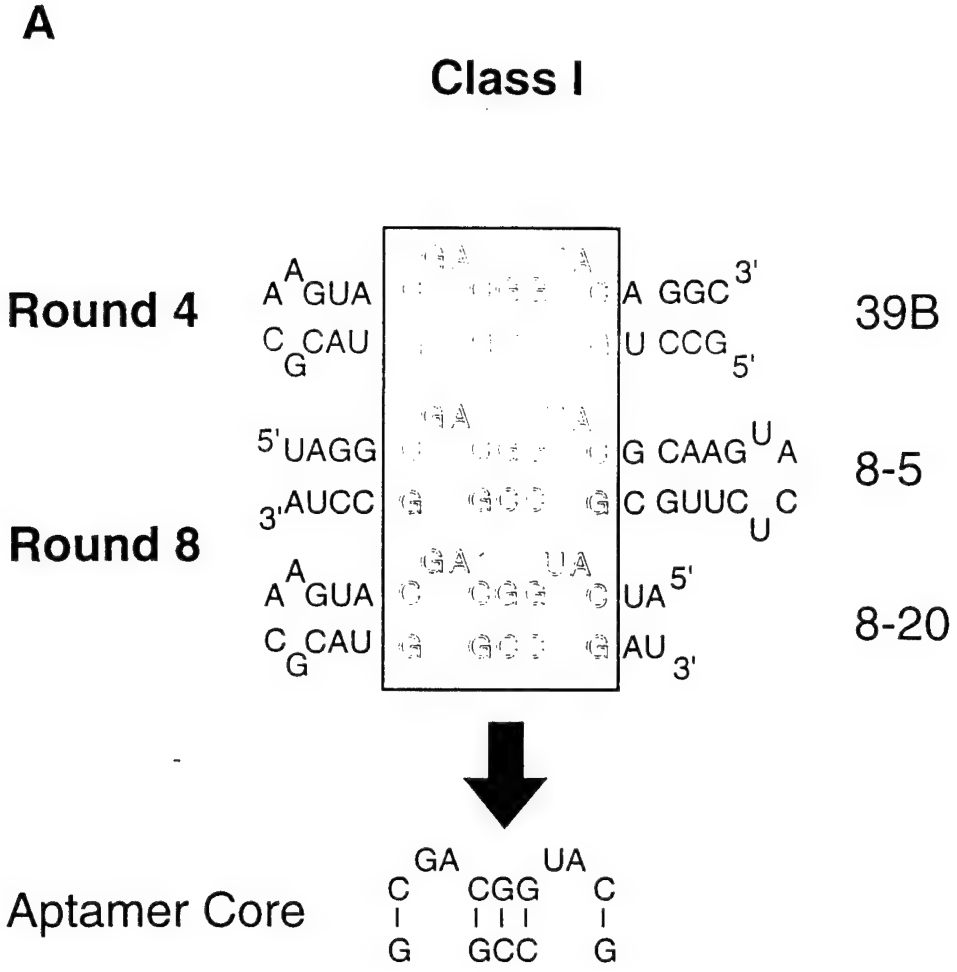


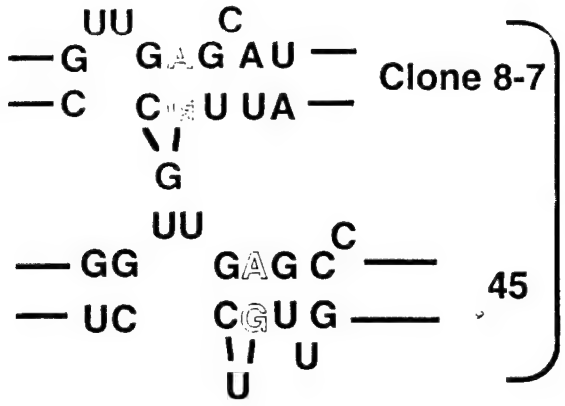
Figure 4. Aptamer motifs from round 4 and round 8 of the 79.9 selection.

(**A**) Class I
The highest affinity sequences from the 79.9 selection fall within the Class I motifs. Comparison of three sequences from round 4 and round 8 of the selected population identified a common aptamer sequence that could be aligned in either orientation relative to the capping stem structure engineered into the pool. Aligning the secondary structures revealed a common set of nucleotides that are assumed to be critical elements for recognition of the Rex protein (boxed region, aptamer core). This element represents a minimal recognition element as identified by *in vitro* selection.

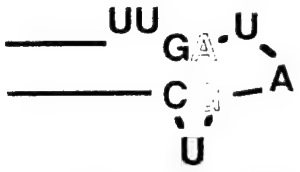
(**B**) Class II
The Class II aptamers are made up of motifs that are very similar to HIV-1 *TAR* elements and Arginine binding motifs that have beem previously identified in other *in vitro* selection experiments (Tao and Frankel, 1996). A set of these contain a proposed homopurine pair that could stabilize an arginine:RNA complex (Tao and Frankel, 1996). Lines indicate sequences outside of the highlighted domain.

Class II

B (HIV-1 *TAR* &Arginine binding motif)



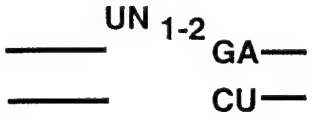
Non-Watson Crick

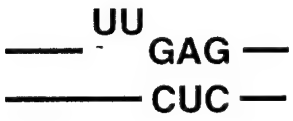


Arginine binding site

(Tao et al., 1996)

TAR (HIV-1)





TAR-like 9D

Characterization of the 8-5 Aptamer

Aptamer 8-5 was assayed for its ability to bind to Rex. A binding curve was generated by incubating increasing amounts of RNA with a constant amount of Rex and capturing aptamer:Rex complexes by filtration. The apparent K_d of the aptamer:Rex complex was found to be 30 nM (Figure 5a). In a complementary experiment, increasing amounts of protein were incubated with a constant amount of RNA and the formation of the aptamer:Rex complex was monitored by a gel mobility shift. The formation of the aptamer:Rex complex was again found to be concentration dependent, and the apparent K_d was similar (25 nM; Figure 5b). The congruence between these different analytical methods indicates that the aptamer is not an artefact of the selection method, and is interacting tightly with the Rex protein.

In order to determine whether the aptamer interacted with Rex at the same site and in the same manner as the *XBE*, competition assays were set up in which a constant amount of radiolabeled stem IID was incubated with increasing amounts of unlabeled aptamer 8-5 in the presence of Rex protein. Complexes were allowed to equilibrate and were then separated from free RNA by filtration. As the aptamer concentration was increased, the fraction of stem IID that was bound was correspondingly reduced. At the highest aptamer concentration tested (5 μM) almost complete (> 90%) inhibition of complex formation was observed, and at an aptamer concentration of only 100 nM 50% inhibition was observed (Figure 6).

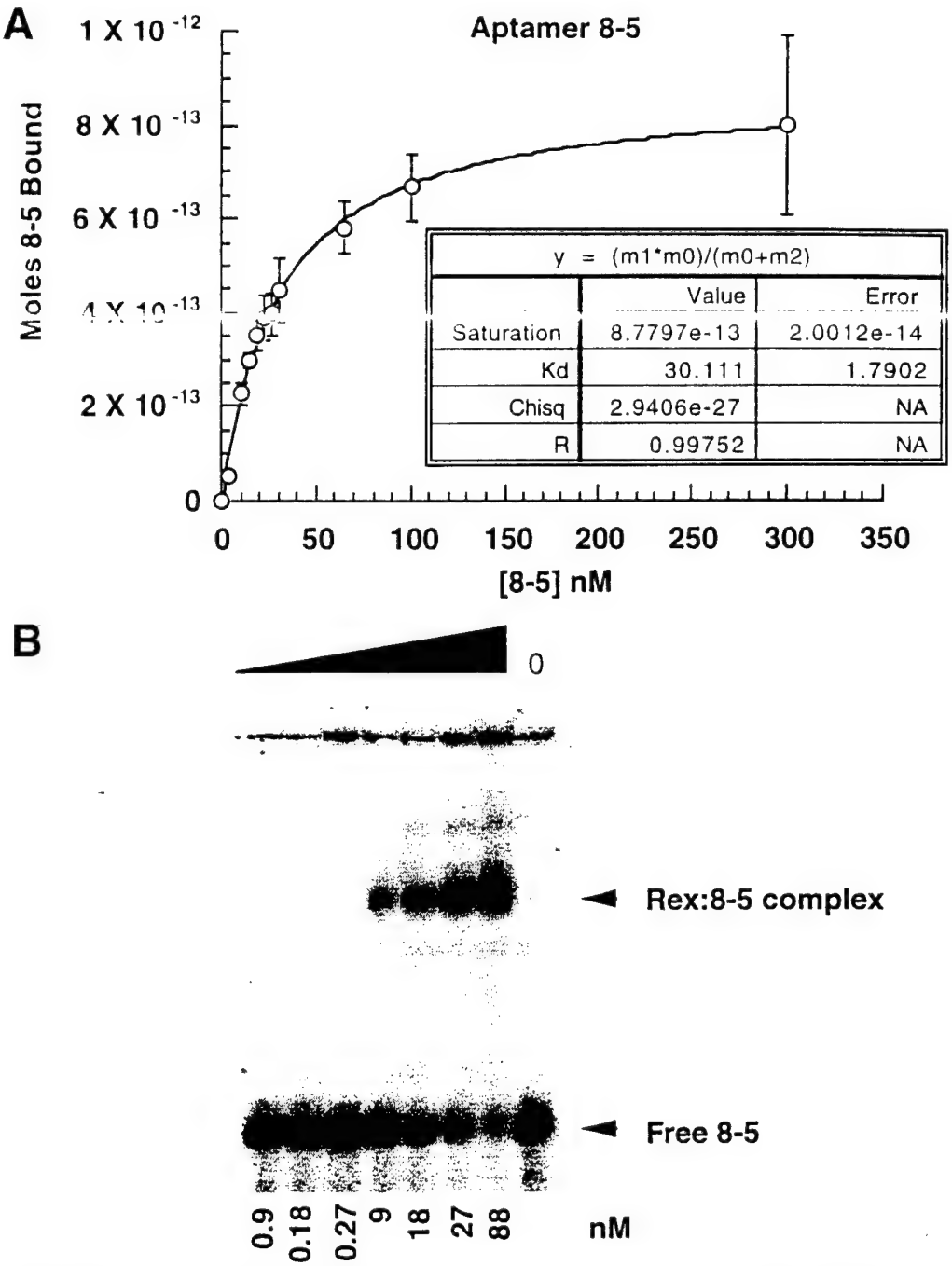


Figure 5. RNA titration and mobility shift of the high affinity aptamer.

(**A**) RNA titration of the high affinity aptamer clone 8-5.
The apparent K_d of the high affinity aptamer (30 nM) was determined by titrating increasing amounts of RNA against a constant amount of protein. K_d's were determined by a least squares fit to the binding data that assumes a simple bimolecular aptamer:Rex interaction using the package Kaleidograph (Abelbeck Software).
(**B**) Mobility shift assay of the high affinity aptamer (8-5).
In mobility shift assays an increasing amount of Rex fusion protein was titrated against a constant amount of body-labeled aptamer RNA (~10 ng). Complexes were separated in 1 X TBE in an 8 %(19:1) native gel at room temperature and visualized using a Phosphorimager. K_d was determined as above.

Inhibition of XRE stem IID Interactions

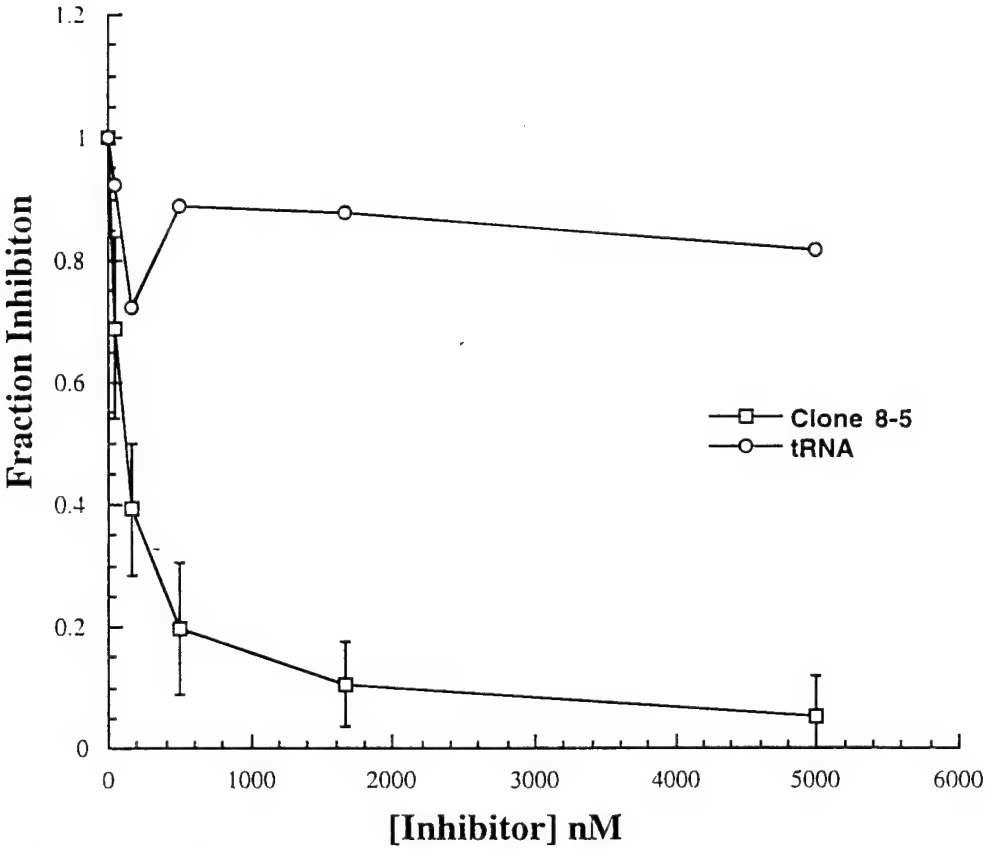


Figure 6. Competition assay of the high affinity aptamer with stem IID of the Rex-responsive element.

Competition for binding to stem IID of the Rex responsive element was measured by incubating increasing amounts of aptamer 8-5 in the presence of a constant amount of body-labeled stem II D (0.5 μM). As the concentration of unlabeled aptamer increases a corresponding decrease in the amount of stem IID retained on the nitrocellulose filter was observed indicating that the aptamer can effectively compete with stem II D for binding to the Rex protein. A 50 % inhibition of complex formation was observed at an RNA concentration of 100 nM determined by a least squares fit to the inhibition data that assumes a simple bimolecular aptamer:Rex interaction using the package Kaleidograph (Abelbeck Software).

The selected anti-Rex aptamer was specific for the Rex protein. For example, when tRNA was tested for its ability to compete with the *XBE* for binding to Rex less than 20% inhibition was observed at concentrations as high as 5 μM. Similarly, in a gel mobility shift assay the anti-Rex aptamer did not form a stable complex with a different ARM protein, HIV-1 Rev (Figure 7a). Moreover, the anti-Rex aptamer could bind to the isolated Rex ARM but not the isolated Rev ARM (Figure 7b). This result is especially significant given that the *XBE* has not been shown to bind to the isolated Rex ARM. The greater affinity of the anti-Rex aptamer for the Rex ARM may have fostered a productive gel-shift.

Aptamers Selected in vitro are Functional in vivo

The high affinity, specificity, and stability of the aptamer:Rex complex and the fact that the aptamer effectively competes with the *XBE* for binding to Rex *in vitro* suggested that the aptamer might effectively mimic the function of the *XBE in* vivo. A reporter system that had previously been used to assay mRNA transport was adapted to the assay of anti-Rex aptamer function (Figure 8a). In the plasmid pCMV138 the chloramphenicol acetyl transferase (CAT) gene has been introduced into a HIV-1 intron, while the *Rev responsive element* has been deleted from this intron. The HIV-1 intron in pCMV138 is spliced out of mRNAs and the gene encoding CAT is degraded. However, when the *XRE* is re-introduced into the HIV-1 intron and the Rex protein is co-expressed, unspliced mRNAs are transported to the cytoplasm and the gene encoding CAT is translated. In effect, the expression of the CAT activity in pDM138 is dependent on the presence of RNA elements and protein factors that can facilitate cytoplasmic transport (Hope et al., 1991;McDonald et al., 1992). This system therefore affords an opportunity to assay aptamers for their ability to substitute for the *XBE* within the *XRE*.

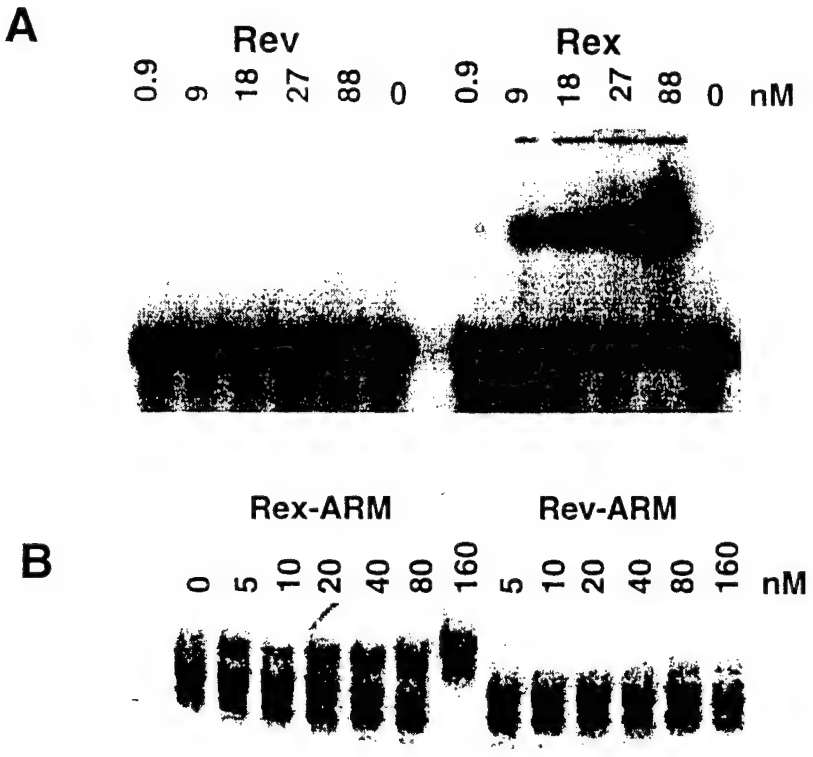


Figure 7. Rex and Rev binding to clone 8-5.

(**A**) Rex and Rev were individually mixed with end-labeled clone 8-5 and allowed to equilibrate for 1 hour at ambient termperature (50 mM KCl; 50 mM Tris-HCl; 2 ug tRNA;~ 10 ng RNA;10 ul total volume). Complexes were separated in a native polyacrylamide gel (8 %;19:1;1 X TBE; 5 watts) at ambient termperature.

(**B**) Rex and Rev ARM binding to min-apt.
Peptides spanning the ARM of Rex and Rev were titrated against a constant amount of end labeled min-apt (50 mM Tris-HCl; 50 mM KCl; 1 mM $MgCl_2$; 5 % glycerol; ~ 5 ng RNA; 10 ul total volume). RNA was incubated with peptide at 4° C for 1 Hour. The complexes were separated on a 10 % native polyacrylamide gel (59:1; .5 X TBE; 350 V) at 4° C.

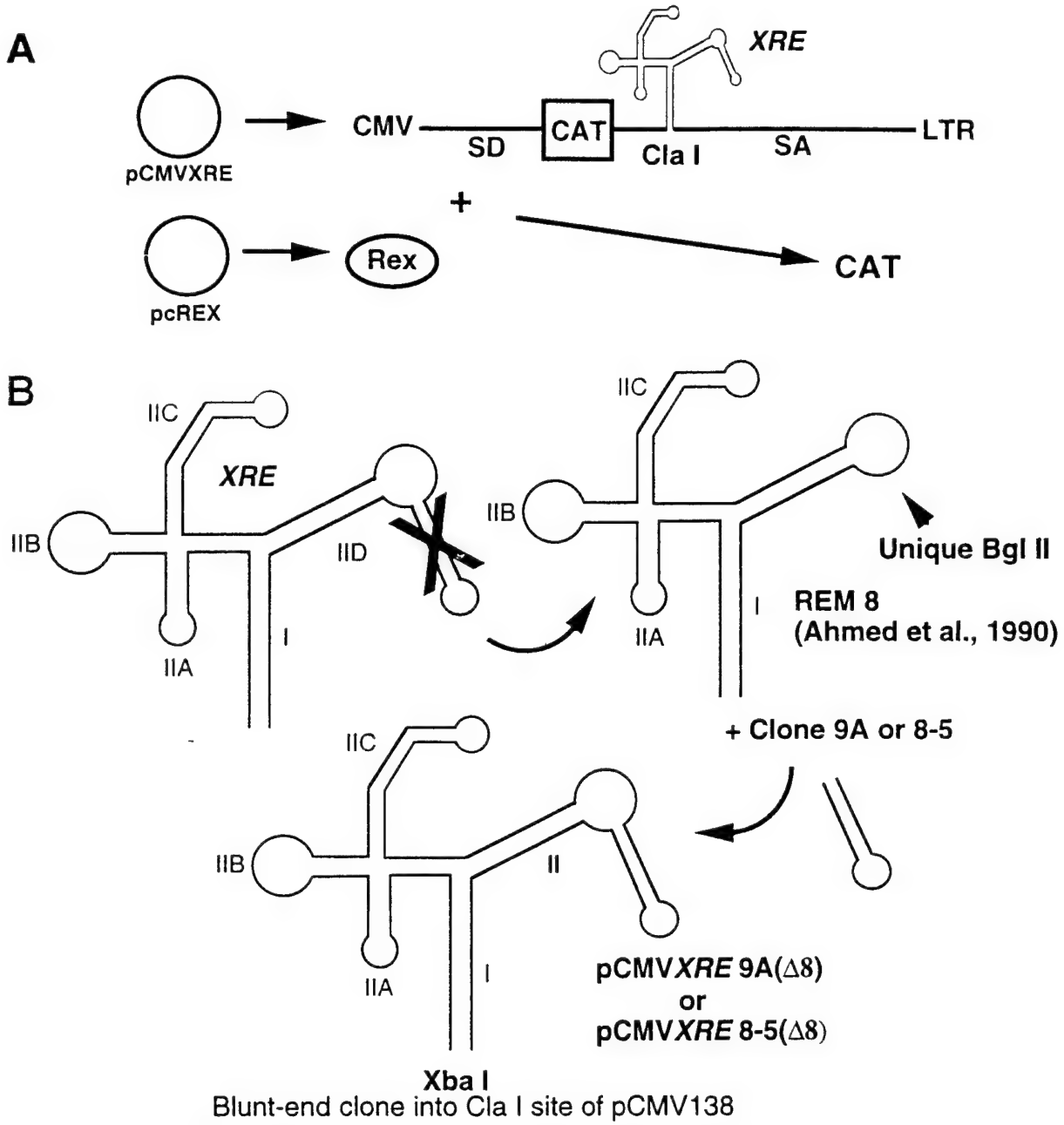


Figure 8. CAT reporter system and chimeric *XRE* constructs.

(**A**) Chimeric XREs were tested for Rex-responsiveness by transiently transfecting CV-1 cells with pcRex (Rex expression plasmid), and a plasmid expressing messenger RNAs containing either wild-type *XRE* or chimeric *XRE* sequences. In the absence of Rex or in the presence of a defective *XRE*, the chloramphenicol acetyl transferase (CAT) gene located within an intron, is completely spliced preventing the production of CAT. A productive *XRE*/Rex interaction promotes the appearance of full length messages containing the CAT reporter gene in the cytoplasm, allowing alterations in the *XRE* to be assayed for function. SD and SA indicate the splice donor and acceptor sites within the construct.

(**B**) Construction of chimeric *XRE*. The Rem 8 deletion (described in Ahmed *et al.*, 1990) was cloned into a pGEM background using unique Xba I sites found at the 5' and 3' end of the *XRE*. Clone sequences (clone 9A or 8-5) were then inserted into a unique Bgl II site which replaces an essential portion of stem II D with the aptamer sequence. This chimeric element was then ligated into pCMV138 at a Cla I site to produce pCMV8-5 (Δ8) and pCMV9A(Δ8).

To this end, a set of hybrid *XRE* sequences were generated in which the XBE was completely replaced with aptamer sequences (Figure 8b). A deletion variant (REM 8, Ahmed et al., 1990) of the *XRE* which deletes a portion of the *XRE* essential for Rex responsiveness , was used to construct a pair of chimeric *XRE* elements that incorporate aptamer sequences in place of the wild-type structures of the *XRE*. Two different aptamer sequences were introduced into REM 8 at the Bgl II site: the high affinity aptamer that was isolated from round 8 (clone 8-5), and an aptamer that was isolated from round 4 but could not effectively compete for binding to stem IID (clone 9A). The low affinity aptamer in effect serves as a negative control for the function of the high affinity site. Plasmids bearing the hybrid *XRE*s were transfected into tissue culture cells along with a Rex expression plasmid, and the Rex-dependent production of CAT activity was determined. The hybrid *XRE* containing the 9A aptamer (Rex 9A (Δ8)) was not Rex-responsive. In contrast the 8-5 aptamer (Rex 8-5 (Δ8)) appeared to efficiently support Rex-responsive mRNA transport, although the hybrid *XRE* was less Rex-responsive than the wild-type *XRE* (Figure 9). The lack of correlation between *in vitro* binding affinity and *in vivo* Rex-responsiveness may be due to a variety of factors, including differences in binding affinities in the cellular milieu relative to the binding reaction, or a slight displacement of Rex on the aptamer relative to the *XBE* that inhibits multimerization.

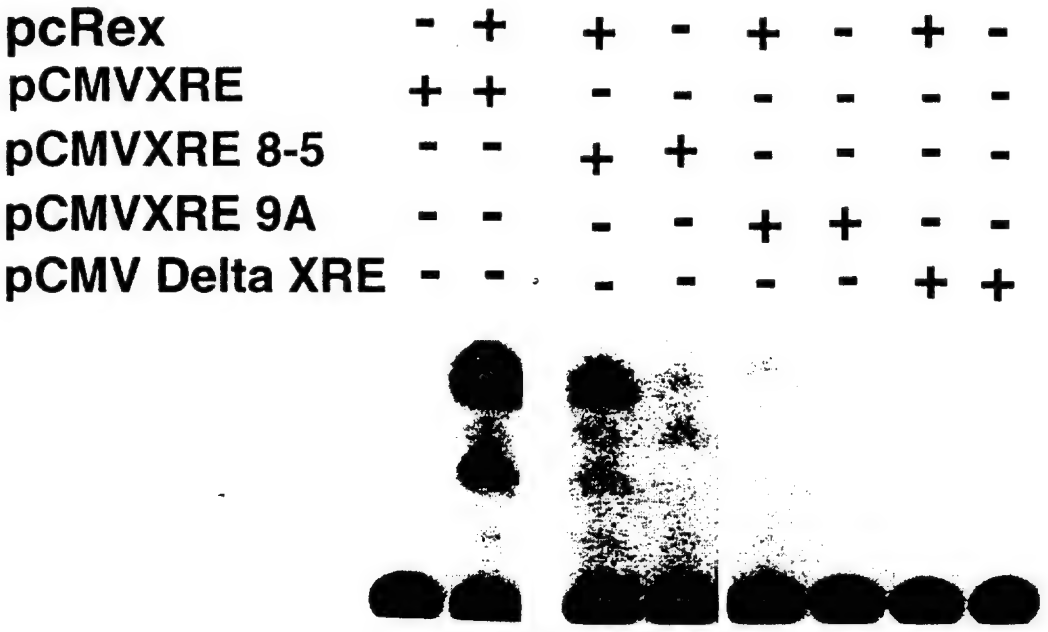


Figure 9. CAT assay of chimeric Rex-responsive elements.

The *XRE* aptamer clone 8-5 or clone 9A was inserted at a unique Bgl II site of a Rem 8 deletion that was not Rex responsive. These aptamer-modified *XRE* fragments were subcloned into the previously characterized CAT reporter plasmid, pCMV138 (Hope, 1991; McDonald et al., 1992). DNAs were co-transfected into CV-1 cells in the presence (+) or absence (-) of the Rex expression plasmid, pcREX (Malim et al., 1989). Forty hours post-transfection, the cells were harvested and assayed for CAT activity as shown. The identity of the selected aptamer incorporated into the *XRE*-containing CAT reporter plasmid is indicated at the left.

DISCUSSION

High-affinity aptamers that bind to a Rex fusion protein of HTLV-I have been isolated from a conformationally constrained pool of RNA. The stringency of the selection was graded in order to efficiently drive the selection process. Protein concentration, competitor concentration, and equilibrium dynamics were systematically varied, and low affinity or non-specifically binding subsets of the RNA were eliminated within only a few rounds. The effectiveness of the strategy was further highlighted by the narrowing of sequence diversity between rounds four and eight: by round eight, low affinity Class III molecules had been eliminated, only a minor set of Class II sequences remained, and the highest affinity binders in Class I predominated. Moreover, the selected pool showed a marked improvement over the unselected pool, and boundRex several fold better than the wild-type *XBE*. The aggregate improvement in binding ability was born out by the affinities of the individual aptamers. When assayed in direct competition with the wild-type *XBE*, the aptamers bound either as well as or up to 9-fold better than the *XBE*.

It can be argued that the anti-Rex aptamers are optimal small RNA ligands for Rex. The pool contained only ca. 10^{11} possible species, a number that was completely spanned by the 10^{13} molecules introduced into the first round of selection (Giver et al., 1993). Thus, all possible RNA internal loop sequences up to and including length 18 were considered. Further, the fact that the sequences were presented in a constrained structure, held between two constant sequence stem loops, implies that selected motifs should largely participate in interactions with the protein rather than in the establishment of structures necessary for the presentation of residues. In support of this point, anti-Rev aptamers were selected from the same random sequence pool and the selected residues were eventually found to directly contact the Rev protein (Ye et al., 1996). While the structural constraints imposed by the constant regions may have excluded other potential structures, such as pseudoknots, there is no evidence that such structures would interact well with ARM

proteins. Sequences and structures similar to the wild-type *XBE* were contained in the original random sequence pool, and known RNA ligands for other ARM proteins, including Rev, HIV-1 Tat, and BIV-1 Tat, would also have been present.

In fact, some of the selected sequences resembled the *XBE*. Class II aptamers from the fourth and eighth rounds of selection contain a conserved sequence motif, 5' UU GAG ... CUC 3', that can be folded into a short stem adjacent to a bulge loop (Figure 10). We have previously used *in vitro* selection to delimit critical Rex-binding sequences and structures in the wild-type *XBE* (Baskerville et al., 1995). One half of the *XBE* is similar to a Class II aptamer (Figure 10). Moreover, the *XBE* half-sites have been found to bind Rex (Baskerville et al., 1995).

The Class II sequence and structural motif is also identical to a sequence and structural motif found in HIV-1 *TAR*. This motif has been shown to form a small arginine-binding pocket that interacts with one of the arginines in the ARM of HIV-1 Tat (Puglisi et al., 1992). In the Tat:*TAR* complex, the guanidino group of an arginine residue forms a pseudo-Hoogsteen pair with the G:C base pair adjacent to the bulge, while the bulged U residue forms a base-triple with A:U that pulls the negatively charged phosphate backbone into apposition with the positively charged arginine. Although the aptamers have two residues in their bulge loops has while *TAR* has three, a mutational analysis of the *TAR* arginine-binding pocket have revealed that two residue bulge loops are optimal (Weeks et al., 1991).

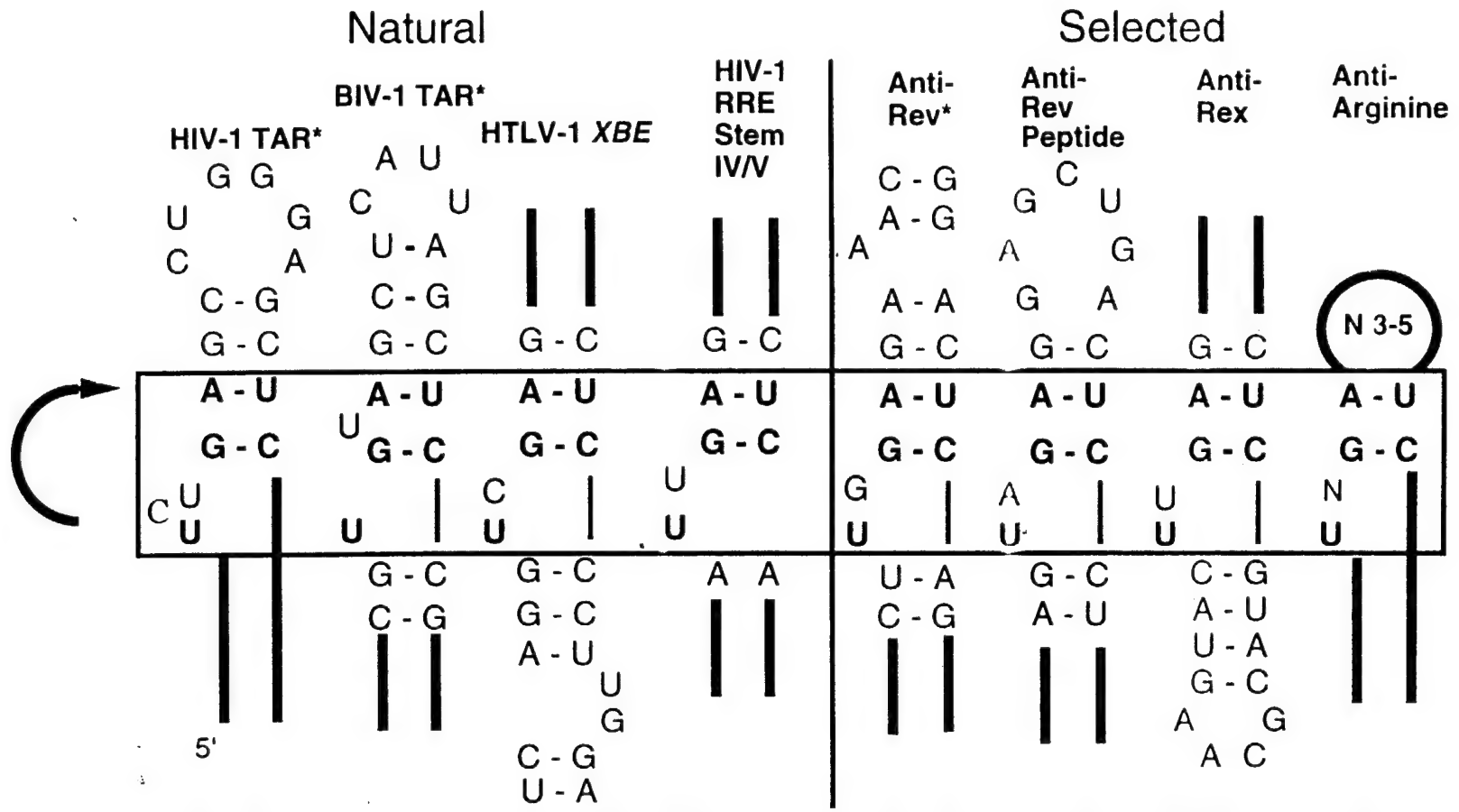


Figure 10. Proposed arginine recognition sites from natural and in vitro selected RNA molecules.

RNA molecules containing a conserved motif that may form a base triple between a bulged uridine and canonical A-U pair in the presence of arginine or an arginine rich motif , (arrow,asterix labeled sequences). HIV-1 *TAR* (Puglisi et al.,1992, Aboul-ela et al., 1995), BIV-1 *TAR* (Ye et al.,1995, Puglisi et al., 1995), and HIV-1 Rev aptamer (Ye et al., 1996), were aligned with motifs proposed to form an arginine binding pocket. The similarity found between these sites and the *in vitro* selected sequences suggests a similar mode of recognition. Lines indicate sequence beyond the displayed region. Anti-Arginine (Tao and Frankel, 1996), Anti-Rev peptide (Xu and Ellington, 1996), Anti-Rev (Giver et al., 1993; Tuerk and McDougal-Waugh, 1993),

It is tempting to speculate that both Class II aptamers and the *XBE* half-site bind to an arginine within the Rex ARM and form *TAR*-like tertiary structures. This hypothesis is bolstered by a wealth of other data. First, sequences that closely resemble the Class II aptamers have been shown to bind arginine *in vitro* (Tao and Frankel, 1996). Second, sequence motifs corresponding to the *TAR* arginine-binding pocket have previously been observed in other natural and selected RNAs that interact with ARM proteins (Giver et al., 1993; Tuerk and MacDougal-Waugh, 1993; Jensen et al., 1994; Xu and Ellington, 1996; Figure 10). The structures of two of these *TAR*-like sequence motifs, those from BIV-1 *TAR* and an anti-Rev aptamer, have been solved and were found to in fact form arginine-binding pockets that resembled the structure of the *TAR* element (Ye et al., 1995; Ye et al., 1996).

The sequences and structures of the Class II anti-Rex aptamers may provide insights into how Rex recognizes other RNA molecules, such as the *Rev-responsive element*. A number of groups have previously found that Rex can recognize the *Rev-responsive element* and functionally substitute for HIV-1 Rev (Rimsky et al., 1988; Hanly et al., 1989; Ahmed et al., 1990; Bogerd et al., 1991; Bohnlein et al., 1991; Unge et al., 1991; Weichselbraun et al., 1992). The primary Rex-binding site on the *RRE* appears to be different from the primary Rev-binding site: Rev binds first to stem II of the *RRE*, while Rex binds to a region that includes stems IV and V (Ahmed et al., 1990; Unge et al., 1991).

The sequence selectivity we have characterized by *in vitro* selection serves as a baseline for understanding Rex recognition of any RNA molecule, and we therefore compared the sequence and structure of the *RRE* with the sequences and structures of the anti-Rex aptamers. Surprisingly, a paired 5' GAG ... CUC 3' triplet is found adjacent to a U-rich single-stranded region. This sequence and structural element present in stem IV/V of the RRE is similar to the arginine-binding pocket found in class II anti-Rex aptamers, anti-Rev aptamers, and other RNA molecules that interact with arginine is (Figure 10).

This putative arginine-binding pocket differs from those seen previously in that a homo-purine base-pair may separate paired and single-stranded regions. However, arginine-binding pockets based on the 5' GAG ... CUC 3' motif have previously been shown to be structurally plastic: the *TAR* element has a three base bulge, mutational analysis of the *TAR* element indicates that two- and four-base bulges are also functional, and BIV-*TAR* splits a portion of the bulge-loop with a base-pair (Figure 10). To the extent that the putative arginine-binding pocket in *RRE* stem IV and class II anti-Rex aptamers are functionally equivalent, our selection results may explain how Rex recognizes the *RRE*.

In contrast, Class I aptamers do not resemble the *XBE*. The primary sequences of the *XBE* and Class I aptamers show no similarities. The *XBE* and Class I aptamers both contain two dinucleotide bulges separated by a central paired triplet; however, the bulged nucleotides are on opposite strands in the *XBE*, but are on the same strand in Class I aptamers. The isolation of non-wild-type elements from completely randomized pools has previously been observed. For example, RNA hairpins selected to bind phage T4 DNA polymerase bore little resemblance to the wild-type sequence (Tuerk and Gold, 1990), while RNA aptamers that bound HIV-1 reverse transcriptase could be folded into a unique pseudoknot structure that differed from a tRNA cloverleaf (Tuerk et al., 1992). Sequence differences between the natural and *in vitro* selected RNAs may be the result of biological restrictions on natural selection that have no counterpart in *in vitro* selection experiments, such as codon usage or RNA folding in the context of a viral genome.

Despite the fact that Class I aptamers do not resemble the *XBE* they may prove to be excellent RNA decoys for Rex. First, our experiments suggest that both the *XBE* and Class I aptamers interact with the Rex ARM. There are conflicting reports in the literature over whether and how the Rex ARM interacts with the *XRE* or *XBE*. On one hand, Hope and co-workers have shown when the Rev and Rex proteins are fused the Rex ARM is dispensible for Rex-dependent mRNA transport (Hope et al., 1990). These results are supported in part by the work of Bohnlein, Hauber, and their co-workers, who have

demonstrated that Rex deletions variants that lack an ARM can nonetheless mediate Rex-dependent mRNA transport, albeit at reduced levels (Hofer et al., 1991). In contrast, Grassman et al. (1991) have shown that Rex deletion variants that lack an ARM do not bind well to the *XRE in vitro*. Non-conservative substitutions in the Rex ARM (Arg5-Arg6-Arg7 to Asp-Leu) disrupt the function of Rex (Rimsky et al., 1989) and the ability of Rex to bind to the *XRE* (Bogerd et al., 1991). Similarly, Hammes and Green (1993) have demonstrated that conservative lysine-for-arginine substitutions within the arginine-rich motif can abrogate Rex-dependent mRNA transport. Since the anti-Rex aptamers bind the Rex ARM and also directly compete with the *XBE* for binding to Rex, our results strongly imply that Rex must also bind to the *XBE* via its ARM. In support of this conclusion, when the Rex ARM is fused to a heterologous proteins (ß-galactosidase), weak binding to the *XRE* can be demonstrated (Grassman et al., 1991). However, as a caveat to this discussion it should be noted that additional *XBE* binding sites within Rex may fall outside the ARM, as suggested by Bohnlein et al. (1991) and Weichselbraun et al. (1992).

Second, Class I aptamers can functionally substitute for the *XBE*. The 8-5 aptamer was cloned into the *XRE* in place of the *XBE* and was found to efficiently support Rex-responsive mRNA transport, though not as well as the wild-type *XRE*. This finding supports and complements other mechanistic studies of viral mRNA transport. Current models for the mechanism of Rex-responsiveness suggest that the cellular mRNA transport apparatus interacts with the Rex protein through an effector domain (Bogerd et al., 1995; Fritz et al., 1995; Stutz et al., 1995). The effector domains of Rex and Rev are similar, and when these domains are swapped for one another unspliced mRNAs appear in the cytoplasm (Weichselbraun et al., 1992). These ARM proteins therefore appear to function as simple connectors that transitively bind viral RNA to the cellular machinery. We now show that the *XRE* may contribute little to this connection beyond the ability to act as a handle for Rex-binding. This interpretation accords with a previous study in which anti-

Rev aptamers substituted for the *RBE* within the *RRE* were found to support Rev-responsiveness (Symensma et al., 1996).

Finally, many of the properties of Class I aptamers can be adapted to the development of reagents for the prevention or treatment of ATL. The anti-Rex aptamer binds specifically to the Rex ARM, eschewing interactions with the similarly charged Rev ARM even at high peptide concentrations, and thus may be unlikely to unintentionally bind to cellular targets. In addition, aptamer 8-5 can be reduced to a 22 nucleotide RNA without loss of binding activity with the incorporation of synthetic linker elements (Osborne S. E., personal communication), allowing its facile synthesis *in vitro*. Similarly, Class I aptamers expressed *in vivo* may serve as RNA decoys that can effectively compete with the *XRE* for binding to Rex and thereby inhibit viral replication. A similar strategy based on the expression of the *RBE* or anti-Rev aptamers has proven successful at disrupting HIV-1 replication (Lee et al., 1994; Good et al., 1997).

REFERENCES

Ahmed Y. F., Hanly S. M., Malim M. H., Cullen B. R., and Greene W. C. (1990). Structure-function analyses of the HTLV-I Rex and the HIV-1 Rev RNA response elements: insights into the mechanism of Rex and Rev action. *Genes & Dev.* **4:**1014-1022.

Bartel D. P., Zapp M. L., Green M. R, and Szostak J. W. (1991). HIV-1 Rev regulation involves recognition of non-Watson-Crick base pairs in viral RNA. *Cell* **67**:529-536.

Baskerville S., Zapp M., and Ellington A.D. (1995). High-resolution mapping of the human T-cell leukemia virus type I Rex-binding element by in vitro selection. *J. Virology* **69**:7559-7569.

Beraud C., Lombard-Platet G., Michal Y., and Jalinot P. (1991). Binding of the HTLV-I Tax1 transactivator to the inducible 21 bp enhancer is mediated by the cellular factor HEB1. *EMBO J.* **10:**3795-3803.

Berkhout B. and van Wamel J. L. (1995). Inhibition of human immunodeficiency virus expression by sense transcripts encoding the retroviral leader RNA. *Antiviral. Res.* **26**: 101-115.

Bogerd H. P., Fridell R. A., Madore S., and Cullen B. R. (1995). Identification of a novel cellular cofactor for the Rev/Rex class of retroviral regulatory proteins. *Cell* **82**:485-494.

Bogerd H. P., Huckaby G. L., Ahmed Y. F., Hanly S. M., and Greene W. C. (1991). The type I human T-cell leukemia virus (HTLV-I) Rex trans-activator binds directly to the HTLV-I Rex and the type 1 human immunodeficency virus Rev RNA response elements. *Proc. Natl. Acad. Sci.* U. S. A. **88:**5704-5708.

Bogerd H. P., Tiley L. S., and Cullen B. R. (1992). Specific binding of the human T-cell leukemia virus type I Rex protein to a short RNA sequence located within the Rex-response element. *J. Virol.* **66**:7572-7575.

Bohnlein E., Berger J., and Hauber J. (1991). Functional mapping of the human immunodeficiency virus type 1 Rev RNA binding domain: new insights into the domain structure of Rev and Rex. *J Virol* **65**:7051-7055.

Cullen B. R. (1992). Mechanism of action of regulatory proteins encoded by complex retroviruses. *Microbiol. Rev.* 375-394.

Ellington A.D. (1989). Purification of oligonucleotides by denaturing polyacrylamide gel electrophoresis. Current Protocols in Molecular Biology, Section 2.12, Ausubel, F.M., et.al. (Eds.), John Wiley and Sons, New York.

Ellington A. D. and Conrad R. (1995). Aptamers as potential nucleic acid pharmaceuticals. *Biotechnology Annual Reveiw* . vol. 1 Edited by El-Gewely MR. *Amsterdam: Elsevier Science.* 185-214.

Fritz C. C., Zapp M. L., and Green M. (1995). A human nucleoporin-like protein that specifically interacts with HIV Rev. *Nature* (London) **376**:530-533.

Gessain A., Barin F., Vernant J. C., Gout O., Maurs L., Calender A., and de The G (1985). Antibodies to human T-lymphotropic virus type-I in patients with tropical spastic paraparesis. *Lancet* **2**: 407-410.

Giver L., Bartel D., Zapp M., Pawul A., Green M., and Ellington A. D. (1993). Selective optimization of the Rev-binding element of HIV-1. *Nucl. Acids Res.* **21**:5509-5518.

Gold L. (1995). Oligonucleotides as research, diagnostic and therapeutic agents. *J. Biol. Chem.* **270**:13581-13584.

Good P. D., Krikos A. J., Li S. X. L., Bertrand E., Lee N. S., Giver L., Ellington A., Zaia J. A., Rossi J. J., and Engelke D. R. (1997). Expression of small, therapeutic RNAs in human cell nuclei. in press, *Gene Therapy.*

Grassmann R., Berchtold S., Aepinus C., Ballaun C., Boehnlein E., and Fleckenstein B. (1991). In vitro binding of human T-cell leukemia virus rex proteins to the rex-response element of viral transcripts. *Virol* **65**:3721-3727.

Gröne M., Koch C., and Grassmann R. (1996). The HTLV-I Rex protein induces nuclear accumulation of unspliced viral RNA by avoiding intron excision and degradation. *Virology* **218**:316-325.

Gröne M., Hoffmann E., Berchtold S., Cullen B. R., and Grassmann R. (1994). A single stem-loop structure within the HTLV-1 rex responsive element is sufficient to mediate rex activity *in vivo*. *Virology* **204**:144-152.

Hammes S. R. and Greene, W. C. (1993). Multiple arginine residues within the basic domain of HTLV-I Rex are required for specific RNA binding and function. *Virology* **193**:41-49.

Hanly S. M., Rimsky L. T., Malim M. H.,. Kim J. H, Hauber J., Duc Dodon M., Le, S.-Y., Maizel J. V, Cullen B. R., and Greene, W. C. (1989). Comparative analysis of the HTLV-I Rex and HIV-1 Rev *trans*-regulatory proteins and their RNA response elements. *Genes Dev.* **3**:1534-1544.

Heus H. A. and Pardi A. (1991). Structural features that give rise to the unusual stability of RNA hairpins containing GNRA loops. *Science* **253**:191-194.

Hidaka M., Inoue J., Yoshida M., and Seiki M. (1988). Post-transcriptional regulator (rex) of HTLV-I initiates expression of viral structural proteins by suppresses expression of rgulatory proteins. *EMBO J.* **7**:519-523.

Hinuma Y., Nagata K., Hanaoka M., Nakai M., Matsumoto T., Kinoshita K. I., Shirakawa S., and Miyoshi I. (1981). Adult T-cell leukemia antigen in an ATL cell line and detection of antibodies to the antigen in human sera. *Proc. Natl. Acad. Sci.* U. S. A. **78**:6476-6480.

Hofer L., Weichselbraun I., Quick S., Farrington G. K., Bohnlein E., and Hauber J. (1991). Mutational analysis of the human T-cell leukemia virus type I trans-acting rex gene product. *J Virol*. **65**:3379-3383.

Hope T. J., McDonald D., Huang X, Low J., and Parslow T. G. (1990). Mutational Analysis of the Human Immunodeficiency Virus Type I Rev Transactivator: Essential Residues near the Amino Terminus. *J. Virol.* 64:5360-5366.

Hope T. J., Bond B., McDonald D., Klein N., and Parslow T. G. (1991). Effector Domains of human immunodeficiency virus type 1 Rev and human T-cell leukemia virus type I Rex are functionally interchangeable and share an essential peptide motif. *J. Virology* **65**:6001-6007.

Inoue J. -L., Yoshida M., and Seiki M. (1987). Transcriptional ($p40^{x}$) and post-transcriptional ($p27^{x-III}$) regulators are required for the expression and replication of human T-cell leukemia virus type I genes. Proc. *Natl. Acad. Sci.* U. S. A. **84**:3653-3657.

Jaeger J. A., Turner D. H., and Zuker M. (1989a). Improved Predictions of Secondary Structures for RNA. *Proc. Natl. Acad. Sci.* USA, Biochemistry. **86**:7706-7710.

Jaeger J. A. ,Turner D. H., and Zuker M.,(1989b). Predicting Optimal and Suboptimal Secondary Structure for RNA in "Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences", R. F. Doolittle ed. *Methods in Enzymology*, **183**:281-306.

Jensen K.B., Green L., MacDougal-Waugh S., and Tuerk, C. (1994). Characterization of an in vitro selected RNA ligand to the HIV-1 rev protein. *J. Mol. Biol.* **235** 237-247.

Klug S. J.and Famulok M. (1994). All you wanted to know about SELEX. *Mol Biol Rep* **20**: 97-107.

Lee S. W., Gallardo H. F., Gilboa E., and Smith C. (1994). Inhibition of human immunodeficiency virus type 1 in human T cells by a potent Rev response element decoy consisting of the 13-nucleotide minimal Rev-binding domain. *J. Virol.* **68**: 8254-8264.

Lin Y. S. and Green M. R. (1989). Similarities between prokaryotic and eukaryotic cyclic AMP-responsive promoter elements. *Nature* **340**:656-659.

Luo Y., Haifeng Y., and Peterlin B. M. (1994). Cellular protein modulates effects of human immunodeficiency virus type 1 Rev. *J. Virol.* **68**:3850-3856.

Malim M. H., Hauber J., Fenrick R., and Cullen B. R., (1988). Immunodeficiency virus rev trans-activator modulates the expression of the viral regulatory genes. *Nature* **335**:181-3.

Malim M. H., Bohnlein S., Hauber J., and Cullen B. R. (1989). Functional dissection of the HIV-1 Rev trans-activator--derivation of a trans-dominant repressor of Rev function. *Cell* **58**:205-214.

Marriott S. J., Lindholm P. F, Brown K. M., Gitlin S. D,. Duball J. F, Radonovich M. F., and Grady J. N. (1990). A 36-kilodalton cellular transcription factor mediates an indirect interaction of human T-cell leukemia/lymphoma virus type 1 TAX_1 with a responsive element in the viral long terminal repeat. *Mol. Cell. Biol.* **10**:4192-4201.

McDonald D., Hope T. J., and Parslow T. G. (1992). Posttranscriptional regulation by the human immunodeficiency virus type 1 Rev and the human T-cell leukemia virus type I Rex proteins through a heterologous RNA binding site. *J. Virology* **66**:7232-7238.

Murphy E. L., Hanchard B., Figueroa J. P., Gibbs W. N., Lofters W. S., Campbell M. Goedert J. J., and Blattner W. A. (1989). Modelling the risk of adult T-cell leukemia/lymphoma in persons infected with human T-lymphotropic virus type I. *Int. J. Cancer* **43**:250-253.

Osame M., Usuku K., Izumo S., Ijichi N., Amitani H., Igata A., Matsumoto M. and, Tara M. (1986). HTLV-I associated myelopathy, a new clinical entity Lancet. **1**: 1031-1032.

Puglisi J. D., Ruoying T., Calnan B. J., Frankel A. D., and Williamson, J. R. (1992). Conformation of the TAR RNA-arginine complex by NMR spectroscopy. *Science* **257**:76-80.

Rimsky L., Hauber J., Dukovich M., Malim M. H., Langlois A., Cullen B. R., and Greene W. C. (1988). Functional replacement of the HIV-1 Rev protein by the HTLV-I Rex protein. *Nature* **335**:738-740.

Rimsky L., Dodon M. D., Dixon E. P., and Greene W. C. (1989). Trans-dominant inactivation of HTLV-I and HIV-1 gene expression by mutation of the HTLV-I Rex transactivator. *Nature* **341**:453-456.

Seiki M., Inoue J., Hidaka M., and Yoshida M. (1988). Two cis-acting elements responsible for posttranscriptional trans-regulation of gene expression of human T-cell leukemia virus type I. *Proc. Natl. Acad. Sci. USA* **85**:7124-7128.

Sullenger B. A., Gallardo H. F. , Ungers G. E., and Gilboa E. (1990). Overexpression of TAR sequences renders cells resistant to human immunodeficiency virus replication. *Cell* **63**: 601-608.

Stutz F., Neville M., and Rosbash M. (1995). Indentification of a novel nuclear pore-associated protein as a functional target of the HIV-1 Rev protein in yeast. *Cell* **82**:495-506.

Symensma T., Giver L., Zapp M., Takle G. B.,and Ellington A. D. (1996). RNA aptamers selected to bind human immunodeficency virus type 1 Rev in vitro are Rev responsive in vivo. *J. Virology* **70**:179-187.

Tan R., Chen L., Buettner J. A., Hudson D., and Frankel A. D. (1993). RNA recognition by an isolated alpha helix. *Cell.* **73**:1031-1040.

Tao J. and Frankel A. D. (1996). Arginine-binding RNAs resembling TAR identified by in vitro selection. *Biochemistry* **35**:2229-2238.

Toyoshima H., Itoh M., Inoue J., Seiki M., Takaku F., and Yoshida, M. (1990). Secondary structure of the human T-cell leukemia virus type 1 rex-responsive element is essential for rex regulation of RNA processing and transport of unspliced RNAs. *J. Virol.* **64**:2825-2832.

Tuerk C. and Gold L. (1990). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249**: 505-510.

Tuerk C., MacDougal S., and Gold L. (1992). RNA pseudoknots that inhibit human immunodeficiency virus type 1 reverse transcriptase. *Proc. Natl. Acad Sci.* U. S. A. **89**: 6988-6992.

Tuerk C. and MacDougal-Waugh S. (1993). *In vitro* evolution of functional nucleic acids: high-affinity RNA ligands of HIV-1 proteins. *Gene* **137**:33-39.

Unge T., Solomin L., Mellini M., Derse D., Felber B. K., and Pavlakis G. N. (1991). The Rex regulatory protein of human T-cell lymphotropic virus type I binds specifically to its target site within the viral RNA. *Proc. Natl. Aca. Sci.* U. S. A. **88**:7145-7149.

Weeks K. M. and Crothers D.M. (1991). RNA recognition by Tat-derived peptides: interactions in the major groove? *Cell* **66**:577-588.

Weichselbraun I., Farrington G. K. , Rusche J. R., Bohnlein E., and Hauber J. (1992). Definition of the human immunodeficiency virus type 1 Rev and human T-cell leukemia virus type I Rex protein activation domain by functional exchange. *J Virol* **66**: 2583-2587

Xu W. and Ellington A. D. (1996). Anti-peptide aptamers recognize amino acid sequence and bind a protein epitope. *Proc. Natl. Acad. Sci.* **93**:7475-7480.

Ye X., Kumar R. A., and Patel, D. J. (1995). Molecular recognition in the bovine immunodeficiency virus tat peptide-Tar RNA complex. *Chem. Biol.* **2**:827-840.

Ye X., Gorin A., Ellington, A. D., and Patel, D. J. (1996). Deep penetration of an alpha-helix into a widened RNA major groove in the HIV-1 rev peptide-RNA aptamer complex. *Nat. Struct. Biol.* **3**:1026-1033.

Yoshida M., Miyoshi. I., and Hinuma, Y. (1982). Isolation and characterization of retrovirus from cell lines of human adult T-cell leukemia and its implication in the disease. *Proc. Natl. Acad. Sci.* U. S. A. **79**:2031-2035.

Zhao L. J., and Giam, C. Z. (1991). Interaction of the human T-cell lymphotrophic virus type I (HTLV-I) transcriptional activator Tax with cellular factors that bind specifically to the 21-base pair repeats in the HTLV-I enhancer. *Proc. Natl. Acad. Sci.* U. S. A. **88**:11445-11499.

Zuker M. (1989). On Finding All Suboptimal Foldings of an RNA Molecule. *Science* **244**:48-52.

Nucleic Acids Symposium Series No. 33 156–159

A simple code for protein:RNA interactions

Andrew D. Ellington, Tonia L. Symensma, Lori Giver and Scott Baskerville
Department of Chemistry, Indiana University, Bloomington, IN 47405, USA

ABSTRACT

The Tat and Rev proteins of HIV-1 and the Rex protein of HTLV-I do not interact with their cognate ligands via a particular structural motif but instead specifically recognize RNA molecules by using agglomerations of arginine residues (1). These proteins are members of the so-called arginine-rich motif (ARM) family. There is little data to support (or contradict) the hypothesis that a few simple arginine:RNA interactions govern how ARMs recognize their viral targets. Not only is it unclear how ARM proteins other than Tat interact with their cognate RNA ligands, for the most part it is not even known how structurally complex these RNA ligands are. In order to fully explore the range of RNA sequences and structures that can bind to ARMs we have carried out *in vitro* genetic selections with two disparate viral proteins: Rev and Rex.

INTRODUCTION

RNA binding proteins fall into a variety of families, most of which have defined amino acid sequence motifs that participate in the recognition of RNA substrates (2). The ARM proteins show no similarity with one another outside of their ARMs, suggesting that this type of protein:RNA interaction may have evolved multiple, independent times. At first glance, this evolutionary independence appears unusual, given that nucleic acid binding proteins within other families appear to arise by modification of an existing structural motif. On closer examination, however, there may be good reasons why the *de novo* accumulation of arginine strings was a viable strategy for some RNA binding proteins. For example, the members of the ARM family are by and large viral proteins that regulate disparate aspects of viral life cycles. As such, these proteins must be able to specifically recognize their targets, generally the viruses themselves, within the context of an extremely high background 'noise:' the multitude of cellular RNA-binding proteins. If a viral protein were to utilize a common cellular RNA-binding motif, such as those employed by hnRNP or zinc-finger proteins, then the possibility of inadvertent cross-talk between viral genomes and cellular regulatory circuits would be maximized.

If ARM:RNA recognition has in fact evolved multiple, independent times in viruses, then it follows that ARM:RNA interactions should be relatively easy to generate and evolve, and, hence, that they should be parsimonious in terms of sequence and simple in terms of structure. From the little that is known about the structure of ARMs and their ligands, this appears to be the case. The paradigm for ARM:RNA recognition is the Tat protein, which specifically binds to a short RNA stem-bulge structure (the *TAR* element) embedded within HIV-1 RNAs (3). Based on NMR data (4), when arginine binds to *TAR* (Figure 1) a three base bulge-loop reorients so that one of the bulged bases forms a U-A:U triple-base interaction and two negatively charged phosphates surround the positively charged guanidino head group of the arginine. At the same time, the O6 and N7 moieties of a guanosine in an adjacent G:C pairing hydrogen bond to the ε and η nitrogens of the arginine.

IN VITRO SELECTION OF APTAMERS

In vitro selection of RNA ligands for Rev

The Rev protein of HIV-1 regulates the splicing and transport of viral mRNAs (5) by binding to a specific site within viral mRNAs, the *Rev responsive element* (*RRE*). The interactions between Rev and the *RRE* can be reduced to a relatively small ARM:RNA interface. Positions 34-50 on the Rev protein are thought to form an alpha helix, and these residues can specifically bind to the *RRE in vitro* (6). Conversely, although functional *RRE*s can span 230 bases, the primary Rev binding site, known as the *Rev binding element* (*RBE*), is much more compact. The *RBE* has been mapped by mutation (7), *in vitro* selection (8), and chemical interference studies (9) to a minimal 30 base stem-internal loop-stem secondary structure on *RRE* stem IIB.

The *RBE* was used as a starting point for selections designed to further define the range of sequences and structures that could specifically bind to Rev (10). An RNA pool was generated (79.9) which contained 12 to 18 randomized positions that encompassed both the internal loop and immediately adjacent flanking sequences; the remaining stems did not resemble the wild-type element. After only three cycles of selection and amplification the pool bound roughly three-fold better than the wild-type *RRE*. Individuals from the fourth cycle were cloned and sequenced.

Sequence analysis revealed that most selected molecules fell into only a few sequence classes (Figure 2). Three major classes of 79.9 aptamers were observed following the selection; the highest affinity aptamers were found in Class I and bound up to ten-fold better than the wild-type element.

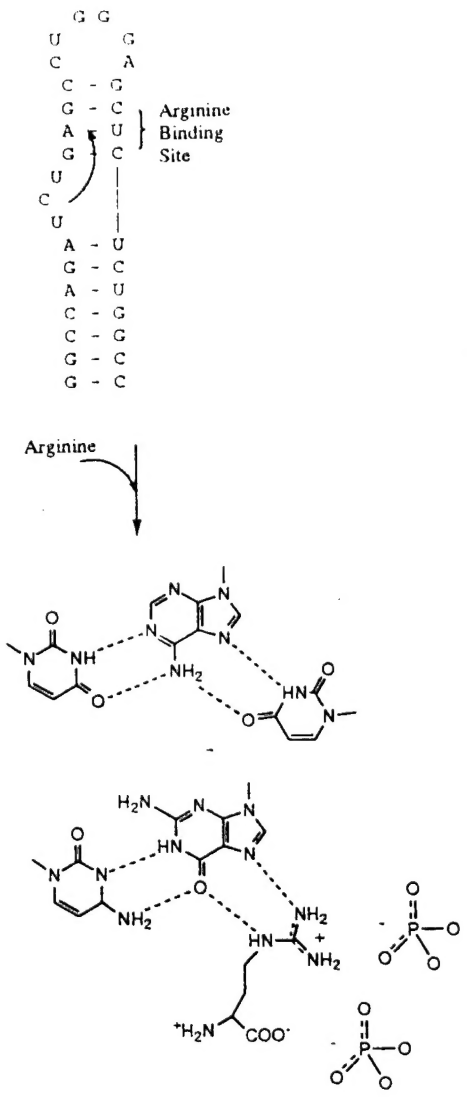


Figure 1. *TAR* undergoes conformational changes on binding arginine.
Upon Tat binding to TAR, the reorientation of three bases results in a U-A:U base triple while the guanidino head group of the arginine binds to the G:C base pair (4).

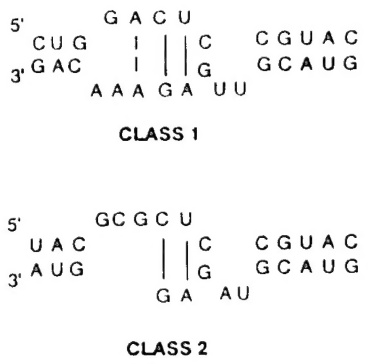


Figure 2. Selected Rev ligands. Examples of two classes of aptamers from the 79.9 Rev selection (10).

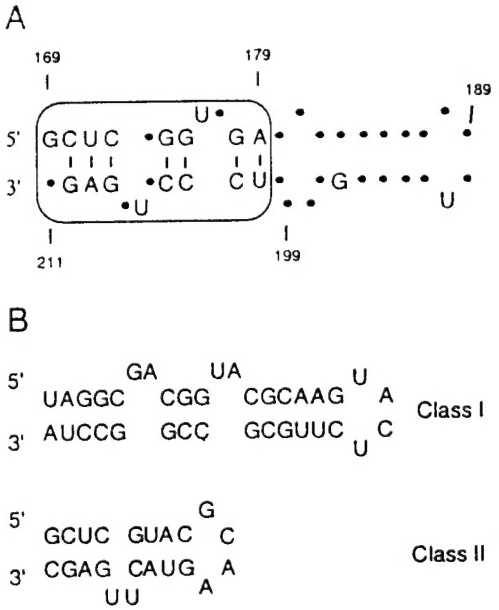


Figure 3. Selected Rex ligands.
(a) Summary of the doped selection. Conserved pairings and single nucleotides are shown. The boxed region spans the primary Rex binding element of the XRE.
(b) Examples of two classes of aptamers from the 79.9 Rex selection.

Class I aptamers appeared to be a chimera of the wild-type sequence and a novel binding motif. The 'left' half of the aptamer resembled the wild-type: non-Watson-Crick pairings could be formed between positions 48 and 71, functionally important residues were present at G47, C49, G70, and A73, and a 'bulged' residue was generally present at position 72. In the 'right' half of the aptamer, however, a new sequence motif 5' CUC ... GAG 3' was found at positions 49-51 and 48-70, respectively, in most (21/31) of the aptamers that were sequenced. Because these sequences could form canonical Watson-Crick base pairs, this motif was assumed to fold into a short helical stack. In addition, this central helix was generally flanked on the 'right' by another novel feature: a bulge of one to three residues, one of which was almost always a uridine. The other two classes of aptamers were variations on Class I. Class II sequences were similar to the 'right' half of Class I; that is, they contained the 5' CUC ... GAG 3' motif and bulged pyrimidines, but were otherwise dissimilar to Class I (and the wild-type sequence).

In order to determine if molecules selected *in vitro* could also function *in vivo*, the best aptamers from the 79.9 selection were cloned into the Rev-responsive element. In the context of the *RRE*, these aptamers could still bind Rev better

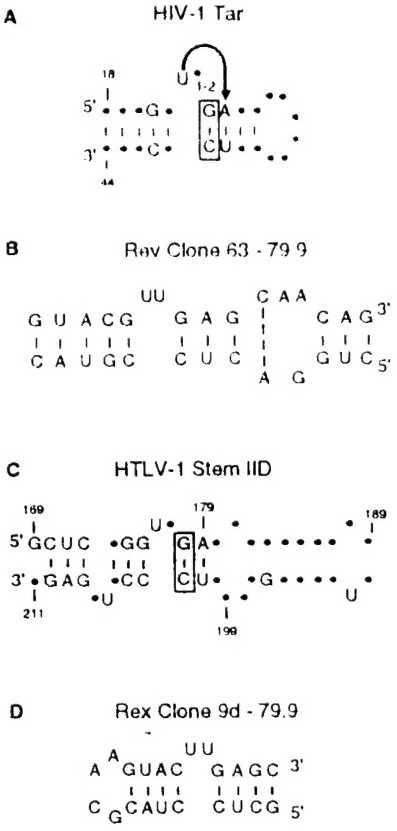


Figure 4. A common arginine-binding motif in ARM ligands.
(a) Sequence and structural features of HIV-1 *TAR*. The structure was redrawn from Weeks and Crothers (15) with numbering according to (4). Residues and secondary structures that contribute to the ability of TAR to bind arginine, a peptide from Tat, or Tat itself are shown. The arrow highlights a triple-base-pair interaction: the guanosine of the boxed G:C pairing interacts with the guanidino group of arginine (4).
(b) A Rev-binding aptamer from the 79.9 Rev selection (10).
(c) HTLV-1 Stem IID structure numbered according to (12).
(d) A Rex-binding aptamer from the 79.9 Rex selection.

than the wild-type sequence, although the effects were attenuated. When the hybrid *RRE*s were transformed into tissue culture cells along with a Rev expression plasmid, they were found to support the transport of mRNAs as well as the wild-type sequence.

In vitro selection of RNA ligands for Rex

The Rex protein of HTLV-I carries out functions analogous to those of Rev in HIV-1 (11). However, these two retroviral proteins are not homologous, and their ARMs bear no similarity to one another. In fact, while the Rev ARM is alpha helical, it is unlikely that the Rex ARM could assume this conformation, as three of the first 16 positions are prolines. Consonant with these facts, Rex recognizes a very different RNA sequence and structure than Rev. The minimal *Rex binding element* (*XBE*), a subsegment of the *Rex responsive element* (*XRRE*), is a partially symmetrical helical motif with bulge loops (11-13).

In order to more completely define the sequences and structures that could be recognized by the ARM of Rex, two random sequence pools centered on the *XBE* were used as starting points for selection. The first was a partially randomized sequence that spanned positions 169 to 211 of the *XBE* and contained 70% wild-type and 30% non-wild-type residues at each position, and the second was 79.9, the pool that was used for the Rev selection. Employing the same pool for both the Rev and Rex selections was intentional, since it allowed the same RNA sequence and structure motifs to potentially be extracted by these different proteins, and, concomitantly, any differences between selected motifs could be interpreted in terms of differences between the ARMs themselves.

While the doped sequence population initially bound Rex much less well than the *XRRE*, after four cycles of selection and amplification the pool could not only compete with the wild-type RNA for binding, but in fact bound 9-fold better. To delimit the Rex binding site, the sequences of selected *XBE*s were aligned and the degree of conservation at each site was determined. The sequences identified by *in vitro* selection to be most important for Rex binding correlated well with the region originally identified by deletion analysis (11).

Results obtained from the doped selection (Figure 3a) precisely defined which residues were most important for Rex recognition by the wild-type *XBE*. In particular, guanosines 169, 174, 175, 178, 196, 208, and 210 were conserved in 50 out of 60 sequences examined ($p<0.05$). Finding multiple, functional G residues within the ARM-binding *XBE* is a first step towards confirming that ARM:RNA interactions may be based on arginine:guanosine interactions. Other conserved residues aided in defining the essential secondary structural features of the *XBE*. For example, cytosines that were predicted to pair with conserved guanosine residues tended to be highly conserved (positions 170, 172, 202, 203, and 204). When base pairs, rather than individual residues, were analyzed, other interactions, such as U171:A209 and A179:U201 were also found to be important for Rex recognition ($p<0.05$).

Prior to the *in vitro* selection results, three different secondary structures had been presented (11,12,14). The information garnered from the doped *XBE* selection was most consistent with the pairing presented by Bogerd *et al.* (12), in which a three nucleotide central helix was flanked by opposing dinucleotide bulges. In particular, the pairing between C172 and G208 is supported by three examples of a covariation to G:U. Similarly, G174:C204 covaries with two other base pairs. While there are no covariations that strongly support tertiary structural interactions within the *XBE*, it is interesting that the sequences and structures defined by *in vitro* selection are consistent with a head-to-head dimeric *TAR* (with a two, rather than three, base bulge;15). If true, two specific arginine binding sites in the *XBE* would have been identified, and it should be possible to use this information to model the relative positions of other amino acids in the complex.

The 79.9 selection yielded aptamers whose sequences fell into three different classes; examples of two of these are shown in Figure 3b. Class I elements contain a highly conserved central stem, 5' CGG ... CCG 3', that is frequently flanked on both sides by bulged nucleotides. The highest affinity aptamers are members of this class, and can bind up to 40-fold better than the wild-type *XRRE* stem IID. Class II aptamers contain the paired stem 5' GAG ... CUC 3' flanked by a bulged diuridine. Surprisingly, the Class II Rex-binding aptamers are very similar to the Class II Rev-binding aptamers that were also selected from the 79.9 pool. This congruence is especially interesting since Rex can functionally substitute for Rev (16).

CONCLUSION

The RNA ligands that bind to Rev are generally different from the RNA ligands that bind to Rex: the *RBE* and most of the Rev-binding aptamers conform to a stem-internal loop-stem structure, while the *XBE* and most of the Rex-binding aptamers form stem-bulge loop-stem structures. Nonetheless, some commonalities are apparent: first, RNA ligands of comparable sizes can specifically recognize their cognate ARMs. More importantly though, several different RNAs that can recognize arginine-rich peptides or proteins may contain the same arginine-binding motif (Figure 4). For example, the paradigm for arginine recognition, the *TAR* element, contains the sequence motif UNGA ... UC that can fold into a specific arginine-binding pocket (Figure 4a). This same sequence motif is found in one of the best Rev-binding aptamers (Figure 4b), in the wild-type Rex-binding element (Figure 4c), and in one class of Rex-binding aptamers (Figure 4d). It is currently unknown whether each of these sequence motifs in fact forms the same structure as is found in the *TAR* element. However, the frequent association of the *TAR* arginine-binding motif with RNA ligands that can specifically bind different arginine-rich proteins suggests that there may be a 'code' for nucleic acid:protein interactions. The basis of this code is the *TAR* arginine-binding motif; additional interactions outside of or adjacent to this arginine-binding motif may expand the range of amino acid sequences that can be bound. If true, this conjecture may allow the design and construction of sequence-specific RNA ligands. In addition, it would provide a ready explanation for our initial observations: viruses have chosen ARM:RNA interactions to regulate their replication because of the simple but modular recognition code that is inherent in such an interface.

ACKNOWLEDGEMENTS

This work was supported by a National Science Foundation National Young Investigator Award (A. Ellington), a Scholar Award from The American Foundation for AIDS Research (A. Ellington), and the Pew Scholar Award in the Biomedical Sciences (A. Ellington).

REFERENCES

1. Mattaj, I.W. (1993) *Cell*, **73**, 837-840.
2. Burd, C.G. and Dreyfuss, G. (1994) *Science*, **265**, 615-621.
3. Cullen, B.R. (1990) *Cell*, **63**, 655-657.
4. Puglisi, J.D., Tan, R., Calnan, B.J., Frankel, A.D., and Williamson, J.R. (1992) *Science*, **257**, 76-80.
5. Cullen, B.R. and Malim, M.H. (1991) *Trends in Biochem. Sci.*, **16**, 346-350.
6. Tan, R., Chen, L., Buettner, J.A., Hudson, D., and Frankel, A.D. (1993) *Cell*, **73**, 1031-1040.
7. Holland, S.M., Chavez, M., Gerstberger, S., and Venkatesan, S. (1992) *J. Virol.*, **66**, 3699-3706.
8. Bartel, D.P., Zapp, M.L., Green, M.R., and Szostak, J.W. (1991) *Cell*, **67**, 529-536.
9. Kjems, J., Calnan, B.J., Frankel, A.D., and Sharp, P.A. (1992) *EMBO J.*, **11**, 1119-1129.
10. Giver, L., Bartel, D.P., Zapp, M., Pawul, A., Green, M., and Ellington, A.D. (1993) *Nucl. Acids Res.*, **21**, 5509-5516.
11. Ahmed, Y.F., Hanly, S.M., Malim, M.H., Cullen, B.R., and Greene, W.C. (1990) *Genes and Development*, **4**, 1014-1022.
12. Bogerd, H.P., Tiley, L.S., and Cullen, B.R. (1992) *J. Virol.*, **66**, 7572-7575.
13. Grone, M., Hoffman, E., Berchtold, S., Cullen, B.R., and Grassman, R. (1994) *Virology*, **204**, 144-152.
14. Toyoshima, H., Itoh, M., Inoue, J.-I., Seiki, M., Takaku, F., and Yoshida, M. (1990) *J. Virol.*, **64**, 2825-2832.
15. Weeks, K.M. and Crothers, D.M. (1991) *Cell*, **66**, 577-588.
16. Rimsky, L., Hauber, J., Dukovich, M., Malim, M.H., Langlois, A., Cullen, B.R., and Greene, W.C. (1988) *Nature*, **335**, 738-740.